



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/00	A2	(11) International Publication Number: WO 98/39427 (43) International Publication Date: 11 September 1998 (11.09.98)
(21) International Application Number: PCT/US98/04421 (22) International Filing Date: 6 March 1998 (06.03.98) (30) Priority Data: 60/039,839 6 March 1997 (06.03.97) US (71) Applicant: UNIVERSITY OF MASSACHUSETTS [US/US]; 55 Lake Avenue, North, Worcester, MA 01655-0243 (US). (72) Inventors: STEIN, Gary, S.; 11 Rice Street, Shrewsbury, MA 01545 (US). STEIN, Janet, L.; 11 Rice Street, Shrewsbury, MA 01545 (US). NILSSON, Susie; 1/6 Monash Street, South Oakleigh 3167, Melbourne, VIC (AU). PETERS, Stefan; Klein Flottbekerweg 74, D-22605 Hamburg (DE). FRENKEL, Baruch; 1620 S. Adams Street, Glendale, CA 91205 (US). HOU, Zhen; Apartment 57, 63 Frank Street, Worcester, MA 01604 (US). QUESENBERRY, Peter; 20 Saxon Lane, Shrewsbury, MA 01545 (US). LIAN, Jane, B.; 96 Carey Avenue, Milton, MA 02186 (US). (74) Agents: MANDRAGOURAS, Amy, E. et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).		(81) Designated States: AU, CA, JP, KR, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: GENE THERAPY USING BONE MARROW TRANSPLANTS TRANSFECTED WITH THERAPEUTIC GENES UNDER THE CONTROL OF TISSUE-SPECIFIC PROMOTERS (57) Abstract <p>Methods for expressing exogenous genes in differentiated cells of a specific type, such as skeletal tissue cells are disclosed. According to the method, pluripotent stem cells capable of maturing into differentiated cells of a specific type are contacted with a nucleic acid comprising an exogenous gene operatively linked to a regulatory element capable of controlling expression of the gene in the differentiated cell. As a result of the method, a population of transduced stem cells capable of maturing into differentiated cells expressing the exogenous gene is produced.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakistan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

GENE THERAPY USING BONE MARROW TRANSPLANTS TRANSFECTED WITH THERAPEUTIC GENES UNDER THE CONTROL OF TISSUE-SPECIFIC PROMOTERS

5 Background of the Invention

Regulatory mechanisms that support osteoblast differentiation and maintenance of bone cell phenotype provide a basis for understanding the structural and functional integrity of the tissue that is principally responsible for skeletal structure and calcium homeostasis. The requirement for continuous renewal of bone, through the remodeling
10 process involving resorption and formation, requires osteoblast proliferation and differentiation throughout the life of an organism.

It is becoming increasingly evident that osteoblast differentiation is a multistep series of events modulated by an integrated cascade of gene expression that initially supports proliferation and the sequential expression of genes associated with the
15 biosynthesis, organizations, and mineralization of the bone extracellular matrix. Equally significant is the growing appreciation of growth factor and steroid hormone-responsive regulatory signals that mediate developmental competency for expression of genes associated with osteoblast proliferation and differentiation. Transcriptional control defines regulatory events operative both developmentally and for support of bone tissue-
20 specific properties.

Over the past several years, studies carried out both *in vivo* and *in vitro* using bone marrow-derived cells have increased our understanding of the processes by which an undifferentiated mesenchymal stem cells undergo commitment and differentiation into osteoblast lineages. These progenitor cells progress through a series of intermediate
25 stages of the osteoblast lineage (Ashton et al. (1984) *Calcif. Tissue Int.*, 36:83-86; Bennett et al. (1992) *The Biological Mechanisms of Tooth Movement and Craniofacial Adaptation*, pp. 91-96; Friedenstein et al. (1970) *Cell Tissue Kinet.* 3:393-403; Haynesworth et al. (1992) *Bone*, 13:81-88; Nakahara et al. (1990) *Bone* 11:181-188, Owen (1980) *Arthritis Rheum.*, 23:1073-1078; Owen (1988) *J. Cell Sci. Suppl.* 10:63-76; Schoeters et al. (1988) *Cell Tissue Kinet.*, 21:363-374; Vilamitjana-Amedee et al. (1993) *In Vitro Cell Dev. Biol.*, 29:699-707). Antibodies developed by several groups that recognize cell-surface proteins of intermediates in osteoblast lineage have contributed to the characterization of stages in the development of bone cell phenotype from marrow-
30 derived osteochondro progenitor cells (Bruder et al. (1989) *Bone*, 11:359-375; Bruder et al. (1989) *Third International Conference on the Chemistry and Biology of Mineralized Tissue*, pp. 73-79; Bruder et al. (1989) *Connect. Tissue Res.*, 20:73-79; Bruder et al. *Bone*, 11:189-198; Haynesworth et al. (1992) *Bone*, 13:69-80; Nijweide et al. (1986)

Histochemistry, 84:342-347; Turkson et al. (1991) *J. Cell Biol.*, 114:373-384).

Although specific regulatory mechanisms that control developmental transitions in these cells remain to be defined, responsiveness to a series of known physiological mediators of osteoblast differentiation have been demonstrated. These factors include bone morphogenic proteins (BMPs), transforming growth factor- β (TGF- β), parathyroid hormone, estrogen and 1,25-dihydroxyvitamin D₃ [1,25(OH)₂D₃].

By the combined application of Northern blot analysis, in situ hybridization, nuclear run-on transcription, and histochemistry, the sequential expression of cell growth and tissue-specific genes has been mapped during progressive development of the bone cell phenotype (Aubin et al. (1993) *Cellular and Molecular Biology of Bone*, pp. 1-45; Gerstenfeld et al. (1987) *Dev. Biol.*, 122:49-60; Owen et al. (1990) *J. Cell. Physiol.*, 143:420-430; Shalhoub et al. (1989) *Biochemistry*, 28:5318-5322; Stein et al. (1993) *Endocr. Rev.*, 14:424-442; Stein et al. (1990) *FASEB J.*, 4:3111-3123). This temporal sequence of gene expression defines four principal developmental periods. Initially, proliferation supports expansion of the osteoblast cell population and biosynthesis of the type I collagen bone extracellular matrix. At this time, genes required for activation of proliferation (e.g., *c-myc*, *f-fos*, and *c-jun*) and cell cycle progression (e.g., histones, cyclins) are expressed together with expression of genes encoding cell adhesion proteins (e.g., fibronectin), as well as, others associated with regulation of extracellular matrix biosynthesis (e.g., TGF- β , type I collagen) and its interrelationships with the cytoskeleton (e.g., integrins). Following the initial proliferation period, expression of genes associated with the maturation and organization of the bone extracellular matrix are upregulated, which contribute to rendering the extracellular matrix competent for mineralization (e.g., alkaline phosphatase). The third developmental period involves gene expression related to the ordered deposition of hydroxyapatite. Osteopontin and osteocalcin exhibit maximal expression at this time when maturation of bone tissue-like organization is ongoing. A fourth developmental period follows in mature cultures during which time collagenase and type I collagen gene expression are elevated, apoptotic activity occurs, and compensatory proliferative activity is evident (Lynch et al. (1994) *J. Bone Miner. Res.*, 9: Suppl.:S352). Although the biological significance of gene expression during the fourth developmental stage remains to be formally established, it appears to serve an editing/remodeling function for modifications in the bone extracellular matrix that sustain the structural and functional properties of the tissue.

The importance of understanding and developing methods for influencing bone cell commitment and differentiation is underscored by the fact that a number of diseases can affect the skeletal system. For example, such diseases include osteoporosis,

- 3 -

immunosuppression drug-induced osteopenia, osteosarcoma, and a series of non-osseous primary tumors (e.g., breast cancer, prostate cancer) which metastasize to bone. Most of the medical procedures used currently for treating such diseases are invasive surgical procedures or toxic drug treatments. Gene therapy is a non invasive method for treating such diseases. However, effective skeletal gene therapy necessitates targeting expression of potentially therapeutic genes to a patient's bone tissue. Accordingly, it would be beneficial to have a method allowing tissue-specific expression of a gene of interest, e.g., skeletal expression of a therapeutic gene, which does not require performing invasive surgery.

Summary of the Invention

This invention pertains to methods for expressing exogenous genes in differentiated cells of a specific type, such as bone cells. The method involves contacting pluripotent stem cells capable of maturing into differentiated cells with a nucleic acid comprising an exogenous gene linked to a regulatory element capable of controlling expression of the exogenous gene in the differentiated cells. As a result of the method, a population of transduced stem cells capable of maturing into differentiated cells expressing the exogenous gene is produced. Preferably, the differentiated cells are in a tissue of interest, such as bone or cartilage, and the exogenous gene is operably linked to at least one osteocalcin regulatory element, such as the OC box I or the OC box II.

Brief Description of the Drawings

Figure 1 is the nucleotide sequence of the human osteocalcin gene (hOC promoter and coding sequences; SEQ ID NO: 1).

Figure 2A is the nucleotide sequence of the mouse osteocalcin gene (mOC promoter and coding sequences). The initial 661 nucleotides upstream of the mouse osteocalcin B gene (mOC-B) joined to the coding region of the mOC gene are shown (SEQ ID NO: 6). Transcriptional regulatory elements are underlined (TATA, OC Box, VDRE) and the translation initiation ATG is indicated in bold (Rahman et al. (1993) *Endocrinology* 133 (6): 3050-3053).

Figure 2B is the nucleotide sequence of the mouse osteocalcin gene (mOC promoter and coding sequences). The first 941 nucleotides contiguous to the mouse mOC-X coding segment are shown joined to the coding region of the mOC gene (SEQ ID NO: 12). The translation initiation ATG is indicated in bold (Rahman et al. *supra*).

Figure 3 is the nucleotide sequence of the rat osteocalcin gene (rOC promoter and coding sequences; SEQ ID NO: 18).

Figure 4 shows an alignment of the nucleotide sequences (promoter and coding sequences) of the human, murine, and rat osteocalcin gene having SEQ ID NO: 1, SEQ ID NO: 6 and SEQ ID NO: 18, respectively. Dots throughout the sequences are introduced to optimize alignment. RNA coding sequences start at position 687.

5 *Figure 5* represents a schematic diagram of regulatory elements and domains of osteocalcin promoters.

Detailed Description of the Invention

10 The invention pertains to a method for expressing an exogenous gene in a cell of a specific type, such as a differentiated cell of a tissue. The method involves contacting pluripotent stem cells with a nucleic acid comprising an exogenous gene operably linked to a regulatory element capable of controlling the expression of the exogenous gene in cells of a specific type, e.g., bone cells. In a preferred embodiment, the population of pluripotent stem cells comprises precursor cells, e.g., osteoprogenitor cells, which are
15 capable of differentiating into cells of specific types, such as bone or cartilage cells. Preferably, the regulatory element is capable of controlling expression of an exogenous gene in the differentiated cell. For example, the exogenous gene is expressed preferentially in a cell of a specific type, such as a cell in which the regulatory element is functionally active. The invention thus provides a method for obtaining cell-type
20 specific expression of an exogenous gene by transfecting a population of pluripotent stem cells with an expressible form of the exogenous gene.

 According to the method of the present invention, a population of transduced stem cells capable of maturing into differentiated cells expressing the exogenous gene is produced. In one embodiment, the transduced stem cells can be administered to a
25 subject. Preferably, the stem cells will home to a characteristic site in the subject and/or differentiate into specific types of cells. Preferred sources of pluripotent stem cells include bone marrow, cord blood and mobilized peripheral blood.

 In one embodiment of the invention, a population of unpurified stem cells, e.g., unfractionated bone marrow, is transfected with a nucleic acid comprising an exogenous
30 gene operably linked to a regulatory element capable of controlling expression of the gene in a differentiated cell of a specific type. A particular advantage of this embodiment of the invention is that isolation of a specific cell population, e.g., progenitor cells, prior to transfection is not required.

 In another embodiment, a population of purified stem cells, e.g., fractionated
35 bone marrow enriched in osteoprogenitor cells, is transfected with a nucleic acid comprising an exogenous gene operably linked to a regulatory element capable of controlling expression of the gene in a differentiated cell of a specific type. The

- 5 -

purification step can include any method of separation of a desired cell population based on, for example, cell surface antigen expression. Examples of purification techniques that can be used include physical separation, e.g., fractionation, panning, as well as positive or negative selection.

5 In a preferred embodiment, the regulatory element capable of controlling expression of the exogenous gene to which it is operably linked in a differentiated cell is an osteocalcin gene regulatory element. In a particularly preferred embodiment, a population of bone marrow cells comprising pluripotent stem cells obtained from a subject, is transfected with a nucleic acid comprising at least one regulatory element
10 from a osteocalcin gene operably linked to an exogenous gene, such that upon differentiation of the stem cells, the exogenous gene is expressed in specific bone cells, e.g., osteoblasts. The bone marrow cells can be unpurified, or purified to enrich for osteoprogenitor cells. Most preferably, the bone marrow cells are unpurified (i.e., total bone marrow).

15 In yet another embodiment, the transfected population of stem cells (purified and unpurified) is administered, e.g., transplanted, to a subject. After administration, the transfected cells migrate to an appropriate site in the subject and differentiate into bone cells which express the exogenous gene.

A decisive advantage of those embodiments using unpurified stem cells is that
20 isolation of progenitor cells, e.g., osteogenic progenitor cells, is not required. Typically, progenitor cells are present only in low numbers in a subject and are difficult to isolate from a whole marrow cell population. Furthermore, progenitor cells can also differentiate into various types of cells, e.g., fibroblasts, chondrocytes, and adipocytes. Accordingly, a method for obtaining expression of an exogenous gene in a tissue-
25 specific manner without requiring isolation of specific cells or progenitor cells is provided.

Promoters and Regulatory Elements

The term "regulatory element" refers to a nucleic acid comprising a nucleotide
30 sequence which is capable of controlling the expression of a gene to which it is operably linked. The term regulatory element is intended to include a nucleotide sequence which, alone or together with one or more other nucleotide sequences is capable of controlling expression of a gene to which the regulatory sequence is operably linked. The terms "promoter" and "regulatory element" are used interchangeably herein and are intended to
35 include promoters, enhancers, polyadenylation signals and other DNA elements regulating expression of a gene. Such regulatory sequences are known to those skilled in the art and are described in Goeddel, *Gene Expression Technology: Methods in*

- 6 -

Enzymology 185, Academic Press, San Diego, CA (1990). A promoter typically contains several regulatory elements.

The term "5' flanking sequence" is intended to include a nucleotide sequence located 5', i.e., upstream, of the transcription initiation site of a gene. Such 5' flanking sequences encompass one or more promoter domains as defined herein. In one embodiment, the 5' flanking sequence is an osteocalcin 5' flanking sequence having a proximal promoter domain, a distal promoter domain and a far distal promoter domain as shown in Figure 5.

The term "3' flanking sequence" is intended to include a nucleotide sequence located 3', i.e., downstream of the polyadenylation signal of a gene.

"Operably linked" is intended to mean that the nucleotide sequence of a gene of interest, e.g., an exogenous gene, is linked to a regulatory element in a manner which allows expression of the gene of interest in a host cell (or in a cell extract). The term operably linked is intended to include a linkage that allows the regulatory element to control the expression of the gene of interest. In a preferred embodiment, a regulatory element is chemically linked to the gene of interest, such as by natural linkage between two nucleotides. The regulatory element can be linked directly to the gene of interest. Alternatively, the regulatory element can be linked indirectly to the gene of interest, such as by a linker, e.g., a nucleotide sequence. Such a linker can vary in length and may contain nucleotide sequences that are recognized by restriction enzymes.

A regulatory element can include a portion of a promoter to which an RNA polymerase binds and which usually contains a TATA box. Such a regulatory element is responsible for the basal transcription of a gene.

A regulatory element can be an enhancer or portion thereof. Generally, an enhancer is an element which modifies the basal transcription of a gene which it controls. An enhancer element can be located several kilobases away from the gene whose expression it affects and it can be located in any part of a gene, such as in 3' or 5' non coding sequences and in introns. Regulatory elements also include silencers, i.e., DNA elements which will reduce the level of transcription of a gene.

Preferred regulatory elements within the scope of the invention include tissue-specific regulatory elements. The term "tissue-specific regulatory element" is intended to include a regulatory element that controls expression of an exogenous gene to which it is operably linked in a specific tissue or cell type. A tissue as defined herein can be an organ, e.g., kidney, liver, heart or a population of cells making up a structure, such as an epithelium. Tissue- or cell-specific expression of a gene occurs predominantly in the cell or tissue in which the regulatory element is active, such as a differentiated cell (and not in an undifferentiated or precursor cell). For example, a gene that is expressed in an

osteoblast under the control of an osteocalcin regulatory element but not in pre-osteoblast can be said to be expressed in a tissue-specific manner.

A preferred regulatory element of the invention is a cell or tissue-specific regulatory element which controls the expression of a gene to which it is operably linked in a cell or tissue of a specific type, such as a bone cell. Accordingly, in a preferred embodiment of the invention, a regulatory element of the invention will have transcriptional activity predominantly in one type of cell. Transcriptional activity in other types of cells is preferentially low or absent. Also within the scope of the invention are regulatory elements which are transcriptionally active in more than one cell type. According to the one application of the invention, it may be desirable to express an exogenous gene in several cell types making up a tissue, such as a bone tissue. Thus, it may be desirable to obtain expression of an exogenous gene in preosteoblasts, osteoblasts, and osteocytes. This may be desirable if, for example, high levels of production of a specific protein, such as a secreted protein, are needed in the bone tissue. Alternatively, it may be desirable to obtain expression of an exogenous gene in a single cell type. Such situations include those in which it is desirable to modify a specific type of cells, such as for correcting a genetic defect or for cell death if, for example the cell is a cancerous cell. It is preferable in these situations that expression is limited to the target cell.

Tissue-specific regulatory sequences for a variety of tissues, including bone tissues, are known in the art. Accordingly, the invention provides methods for tissue-specific expression of a gene in a variety of tissues by, e.g., *ex vivo* transfecting a population of bone marrow cells with a gene under the control of said tissue-specific regulatory element and administration of the transfected bone marrow cell population to a patient.

Regulatory elements of the invention can be inducible. The term inducible regulatory element is intended to mean that the expression of an exogenous gene that is controlled at least in part by a regulatory sequence can be modified, by for example, exposing the cell to a specific compound or to a specific condition. Inducers within the scope of the invention include agents which interact with a receptor on the surface of a cell or within a cell and include among others hormones, cytokines, chemicals. Regulatory elements which are inducible are well known in the art. In one embodiment, the regulatory element is inducible and tissue-specific. For example, the osteocalcin promoter controls osteoblast specific expression and is inducible by, for example, Vitamin D, steroid hormones, and glucocorticoids. Regulatory elements can also be induced by a change in the environment of the cell or tissue, such as a change in temperature, (e.g., an increase in temperature) and include, for example, nucleotide

sequences termed "heat shock elements". Alternatively, a regulatory element can be induced by exposure to light of a certain wavelength, such as ultraviolet light. In yet another embodiment, a regulatory element is induced by shock. Inducible regulatory systems for use in mammalian cells are known in the art, for example systems in which

5 gene expression is regulated by heavy metal ions (Mayo et al. (1982) *Cell* 29:99-108; Brinster et al. (1982) *Nature* 296:39-42; Searle et al. (1985) *Mol. Cell. Biol.* 5:1480-1489), heat shock (Nouer et al. (1991) in *Heat Shock Response*, e.d. Nouer, L., CRC, Boca Raton, FL, pp167-220; Morris T. (1991) *Mol. Cell. Biology* 11:544), hormones (Lee et al. (1981) *Nature* 294:228-232; Hynes et al. (1981) *Proc. Natl. Acad. Sci. USA*

10 78:2038-2042; Klock et al. (1987) *Nature* 329:734-736; Israel & Kaufman (1989) *Nucl. Acids Res.* 17:2589-2604 and PCT Publication No. WO 93/23431), tetracycline (Gossen, M. and Bujard, H. (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551 and PCT Publication No. WO 94/29442) or FK506 related molecules (PCT Publication No. WO 94/18317)

15 In view of the knowledge in the art relating to tissue-specific and inducible regulatory elements, also within the scope of the invention are regulatory elements which have been modified. For example, a regulatory element that is tissue-specific for bone, e.g., osteocalcin promoter, can be engineered to become inducible by a specific compound by operably linking the tissue-specific regulatory element with a synthetic

20 element responsive to the compound.

In a preferred embodiment of the invention, the regulatory elements are capable of controlling expression of the exogenous gene preferentially in bone cells. The term "bone cells" is intended to include any cell which is present in bone tissue, such as preosteoblasts, osteoblasts, osteocytes, osteoclasts, or precursors of these cells which are

25 localized in a bone tissue. Bone tissue is intended to include trabecular bone, intramembraneous bone, and cancellous bone, among others.

Preferred regulatory elements of the invention include nucleotide sequences derived from an osteocalcin gene. Osteocalcin is a major component of the bone extracellular matrix and is as abundant as collagen on a molar basis. Expression of the

30 osteocalcin gene is highly tissue-specific, being most abundant in skeletal tissue. In particular, expression has been found to occur at highest levels during one period in the osteoblast developmental sequence, as described, e.g., in Stein et al. (1996) *Physiological Reviews* 76:593-629. Briefly, the first period of osteoblast development is associated with proliferation of osteoblasts, expansion of the osteoblast cell population and biosynthesis of the type I collagen bone extracellular matrix. The second period of

35 osteoblast development is associated with the maturation and organization of the bone extracellular matrix to render it competent for mineralization. In the third period of

osteoblast development osteocalcin expression is highest, and this period is associated with the deposition of hydroxyapatite. The fourth developmental period is associated with apoptotic activity and compensatory proliferative activity. Accordingly, use of an osteocalcin regulatory element can result in tissue-specific transcriptional activity of an exogenous gene.

The human osteocalcin gene has been localized to the 1q distal region of chromosome 1 and a mouse osteocalcin gene has been localized to chromosome 3 (Puchacz et al. (1989) *Endocrinology* 24: 2648-2650; Lian et al. (1989) *Proc. Natl. Acad. Sci. USA* 86: 1143-1147). In the mouse, osteocalcin is a multi-gene family which includes three osteocalcin genes. Two of the three murine osteocalcin genes have the same promoter and the third gene has a different promoter that is expressed in several non-bone tissues such as brain, lung, and kidney. The transcribed regions of the osteocalcin genes expressed in bone of a rat, human, and mouse contain four exons and three introns. The promoters of these mammalian osteocalcin genes are very similar, in particular in regard to the type and location of regulatory elements. The nucleotide sequence of the human osteocalcin gene (promoter and coding sequences) is represented in Figure 1 and corresponds to SEQ ID NO: 1. Amino acid sequences corresponding to exon sequences of the human gene are represented in SEQ ID NOs: 2-5. The nucleotide sequence of two portions of the bone-specific mouse osteocalcin gene (promoter and coding sequences) are represented in Figures 2A and 2B and corresponds to SEQ ID NOs: 6 and 12. Amino acid sequences corresponding to exon sequences of the mouse gene are represented in SEQ ID NOs: 7-11 and SEQ ID NOs: 13-17, respectively. The nucleotide sequence of the rat osteocalcin promoter is represented in Figure 3 and corresponds to SEQ ID NO: 18. Amino acid sequences corresponding to exon sequences of the rat gene are represented in SEQ ID NOs: 19-22. Figure 4 represents a sequence comparison of the human osteocalcin promoter and coding sequences (SEQ ID NO: 1), the mouse osteocalcin promoter and coding sequences (SEQ ID NO: 6), and the rat osteocalcin promoter and coding sequences (SEQ ID NO: 18) and indicates the strong similarity between these mammalian promoters.

The structure and function of the rat osteocalcin regulatory element is described, for example, in Stein and Lian (1995) *Endocrine Reviews* 4:290-297 and Stein et al. (1996) *Physiological Reviews* 76:593-629. A schematic diagram of the rat osteocalcin regulatory element and certain promoters is represented in Figure 5. The rat osteocalcin regulatory element is represented as comprising two major promoter domains which are flanked by DNase I hypersensitive sites in osteocalcin-expressing cells (Montecino et al. (1994) *Biochemistry* 33:348), and which encompass cis-acting elements contributing to regulation of promoter activity.

As used herein, the signs "-" and "+" followed by a number refer to the number of nucleotides upstream and downstream, respectively, from a reference site, e.g., the transcription initiation site. The proximal promoter domain of the osteocalcin gene, located from about -0.2 kb upstream from the TATA box (located at about -42 to -39 upstream from the transcription initiation site) controls basal transcriptional level and tissue-specific expression (Lian et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:1143). This domain is mediated primarily by two cis-acting elements, designated OC Box I and OC Box II (reviewed in Stein et al. (1996) *Physiological Reviews* 76:593-629). The OC box I (located at about -0.1 kb upstream from the transcription initiation site) is a highly conserved regulatory element in the mammalian osteocalcin genes required for basal and tissue-specific expression and contains multiple regulatory elements including an AP-1 sequence and homeodomain binding sites (MSX). FOS/JUN-related proteins form heterodimers at the AP-1 site. The OC box I binds homeodomain-containing transcription factors (Hoffmann et al. (1994) *J. Cell. Biochem.* 61:310-324; Hoffman et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:12887; Towler et al. (1994) *Mol. Endocrinology* 8:1484; Towler et al. (1994) *Mol. Endocrinology* 8:614). Competition and mutational analysis and protein binding experiments have shown that the homeodomain proteins Msx-1 and Msx-2 bind the ATTA motif in the OC box I. A bone specific transcription complex designated OC box I binding protein (OCBP) has also been shown to bind the OC box I (Hoffmann et al. (1996) *J. Cell. Biochem.* 61:310-324).

OC Box II (located about -0.149 kb upstream from the transcription initiation site) binds transcription factors of the acute myelogenous leukemia (AML) family of transcription factors, and over expression of the transcription factor AML-1 in non-osteous cells results in activation of a transiently transfected osteocalcin promoter construct (Merriman et al. (1995) *Biochemistry* 34:13125-13132; Banerjee et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:4968-4973). The proximal promoter of osteocalcin genes also contains an E-box that bind the Id helix-loop-helix (HLH) proteins. This sequence is believed to play a role in the suppression of osteocalcin gene expression. Another functional element in the OC promoter is the OC Specific Element 1 (OSE₁) (Ducy et al. (1995) *Mol. Cell. Biol.* 15:1858).

Several elements in the osteocalcin gene control responsiveness to hormones and cytokines. For example, glucocorticoid responsive elements (GREs) are located at about -16, -86, and -697 nucleotides from the site of initiation of transcription (Heinrichs et al. (1993) *Biochemistry* 32:11436-11444; Alsam et al. (1995) *Mol. Endocrinology* 9:679). These sites are bound by glucocorticoid receptors. It has also been reported that the transcription factor NF-IL6 binds these proximal GREs.

It has also been shown that several fold ligand-dependent enhancement of basal promoter activity is mediated by binding of the vitamin D Receptor/Retinoid X Receptor (VDR/RXR) heterodimer to a vitamin D response element (VDRE) located at about -465 nucleotides upstream from the transcription initiation site (Markose et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:1701). The VDRE is similar to the superfamily of related steroid response elements including the estrogen-responsive element, the thyroid hormone-responsive element, GRE, and the retinoic acid-responsive element. The minimal VDRE is composed of two half steroid motifs which are either direct or indirect repeats separated by 3 nucleotides. Tumor necrosis factor- α (TNF- α) and retinoic acid have also been reported to affect the VDRE transcription factor complex which binds the VDRE. TNF- α regulation also involves, at least in some cases, an NF κ B site.

Around nucleotide -144 upstream from the osteocalcin regulatory elements is a TGF β responsive element (TGRE) (Banerjee et al. (1996) *Endocrinology* 137:1991). This element is involved in downregulation of the osteocalcin gene by treatment with TGF- β 1. Deletion analysis of the rat osteocalcin promoter has indicated that the TGF- β 1 responsive element corresponds to a 29-bp region from about -162 to about -134 of the promoter. This element contains an AP-1 site, which has been shown to be involved in regulation of the promoter by TGF- β 1 and which binds Fra-2 and Jun-B proteins.

It has been shown that a negative regulatory domain contributing to developmental expression of the osteocalcin gene resides in the coding region overlapping the first exon (Frenkel et al. (1993) *Biochemistry* 32:13636 and Frenkel et al. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91:10923). This domain is termed the OC silencer and contains a ACCCTCTCT sequence motif, present in several tissue-specific genes. Accordingly, it may be desirable to include this silencer element in the nucleic acid of the invention for obtaining preferentially bone specific expression.

In one embodiment, sequences between about -1.7 and about -1.1 kb are employed for high level activity of the osteocalcin gene when the gene under control is chromosomally integrated (Frenkel et al. *J. Bone Min. Res.* 11 (Supl. 1) S147, Abstract # P225). Indeed, transgenic mice carrying the CAT reporter gene under the control of the osteocalcin -1700/+24 promoter express CAT specifically in skeletal tissues, and at higher levels than the expression observed with shorter promoter sequences (see Example 1). Expression is highest at an early age, but substantial promoter activity is maintained throughout life and vary as a function of skeletal site and sex (Frenkel et al. *supra*). Transgenic mice containing 3.9 kb of the human osteocalcin promoter fused to a CAT reporter gene, express CAT predominantly in bone, but also at reduced levels in the brain, hypertrophic chondrocytes and kidney (Kesterson et al. (1993) *Mol. Endocrinol.* 7:462). Low level expression in tissues other than bone in these transgenic

- 12 -

animals is consistent with low level osteocalcin expression in some tissues other than bone, e.g., bone marrow megakaryocytes and peripheral blood platelets (Thiede et al. (1994) *Endocrinology* 135:929; Barille et al. (1996) *J. Bone and Min. Res.* 11:466; Fleet and Hock (1994) *J. Bone and Min. Res.* 9:1565).

5 Accordingly, in one embodiment of the invention, an exogenous gene is operably linked to at least a portion of a 5' flanking sequence of an osteocalcin gene, such that expression of the exogenous gene occurs preferentially in bone cells. Preferred regulatory elements of the invention include portions of the 5' flanking sequence of an osteocalcin gene (e.g., promoters, enhancers) necessary to obtain bone-specific
10 expression of an exogenous gene operably linked to such a promoter. Preferred osteocalcin regulatory elements include the proximal promoter, distal promoter and far distal promoter. In addition, preferred osteocalcin regulatory elements include the OC box I, the OC box I and the OC box II, or the OC box I and the VDRE of an osteocalcin gene. Also within the scope of the invention are promoters containing additional
15 combinations of regulatory elements of the osteocalcin gene represented in Figure 5, described above, or known in the art, so long as bone specific expression of a gene to which the promoter is operably linked occurs. Other preferred regulatory elements of the invention include combinations of the proximal promoter of an osteocalcin gene, the distal promoter of an osteocalcin gene and the far distal promoter of an osteocalcin gene.

20 Other preferred osteocalcin regulatory elements include such fragments of 5' flanking sequence of an osteocalcin gene as 2.0 kb, 1.5 kb, 1.0 kb, or 0.5 kb. These fragments are preferably fragments from an osteocalcin promoter having a nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO: 6, SEQ ID NO: 12, or SEQ ID NO: 18. Particularly preferred 5' flanking sequence are fragments of an osteocalcin fragments of
25 the human osteocalcin promoter, such as a fragment from SEQ ID NO: 1.

 Also within the scope of the invention are regulatory elements which control tissue-specific expression of a gene and are functional equivalents or fragments or modified forms of the regulatory elements set forth above. Regulatory elements which are homologous to the regulatory elements set forth herein and which are capable of
30 directing tissue-specific expression of a gene are also within the scope of the invention. Such functional equivalents and homologs are intended to include nucleic acid molecules which share sequence similarity to a sequence of SEQ ID NO:1, SEQ ID NO: 6, SEQ ID NO: 12, or SEQ ID NO: 18. Preferred modifications of an osteocalcin regulatory sequence include those affecting nucleotides outside of promoters and
35 enhancers as defined herein. However, in view of the extensive knowledge in the art of osteocalcin regulatory elements, a person of skill in the art would know which portions of such regulatory elements can be modified by nucleotide substitution, addition or

deletion while the desired functionality is maintained. Furthermore, numerous assays can be performed to confirm that a functional equivalent or homolog of an osteocalcin regulatory element is capable of controlling bone-specific expression of a gene to which it is operably linked. Examples of these assays are set forth herein or are known in the art.

Modifications of a natural regulatory elements of a tissue-specific promoter can also be performed to increase or decrease expression, or to change the responsiveness of the regulatory element to an agent. Such modifications can be performed by nucleic acid deletions, substitutions, or additions.

Osteocalcin regulatory elements from species other than human, mouse or rat can also be used in the invention so long as it is capable of appropriately controlling expression of an exogenous gene to which it is operably linked. Osteocalcin regulatory elements from species other than human, mouse or rat can be isolated by methods known in the art, e.g. low stringency hybridizations. A preferred regulatory element for controlling expression of an exogenous gene in bone cells is a fragment of a 5' flanking sequence of the osteocalcin gene including from about nucleotide 1 to about nucleotide 700 of the 5' flanking sequence, such as the human osteocalcin promoter having SEQ ID NO: 1. Another preferred fragment includes from about nucleotide 1 to about nucleotide 350 of a 5' flanking sequence of an osteocalcin gene, such as a fragment of SEQ ID NO: 1. Another preferred fragment includes at least 1.7 kb 5' flanking sequence preferably starting at the initiation of transcription.

In one embodiment, the nucleic acid of the invention comprises at least two regulatory elements, such as two of the same or different promoters. For example, the regulatory element can include a human osteocalcin promoter and a mouse osteocalcin promoter.

Also within the scope of the invention are constructs containing both 5' flanking and 3' flanking sequences from an osteocalcin gene operably linked to an exogenous gene. Such constructs are described, e.g., in the PCT application number WO 96/05299.

Numerous assays can be performed to confirm that a promoter or a regulatory element is capable of controlling the expression of an operably linked exogenous gene in the desired tissue, e.g., bone. For example, a fragment of a regulatory element, such as a fragment of an osteocalcin promoter can be fused to a reporter gene, such as the gene encoding the chloramphenicol transferase (CAT) or luciferase and transfected into a cell line. Preparation of these constructs and introduction into cells can be performed according to standard techniques. For this assay, it is desirable to transfect several cell lines with the construct. Some of these cell lines should correspond to cells in which expression is desired and other cell lines should correspond to cells in which expression

is not desired. For example, to confirm that a fragment of an osteocalcin promoter is active in osteoblasts, this promoter fragment can be linked to the CAT gene and transfected in osteoblast cell lines, which should express the reporter gene and in cells such as muscle cells and adipocytes, which should not express the reporter gene.

5 Another assay for determining the activity of a regulatory element is transfection of the regulatory element operably linked to a reporter gene into a cell line that is capable of differentiating *in vivo*. In this assay, the cell differentiates into a cell in which expression of the exogenous gene is desired. For example, osteoblast progenitors from bone marrow are pluripotent cells that can differentiate *in vitro* into fibroblasts,
10 chondrocytes, and adipocytes (Beresford et al. (1993) *Am. J. Med. Genetics* 45:163-178). Mouse, rat and human bone marrow-derived cells maintained under protocols that employ serum, ascorbic acid, β -glycerophosphate and dexamethasone, express bone phenotypic molecular markers and form mineralized nodules (Maniatopoulos et al. (1988) *Cell & Tissue Res.* 254(2):317-330; Beresford et al. (1993) *Am. J. Med. Genetics*
15 45:163-178; Cheng et al. (1994) *Endocrinology* 134:277-286; Kassem et al. (1994) *Calcif. Tissue Int.* 54:1-6; van Auken and Baran (1994) *J. Bone & Min. Res.* 9 (Supp. 1) A:26). When these marrow cells contain a reporter gene under the control of the bone specific osteocalcin promoter, the reporter gene is expressed concomitant with the differentiation of the marrow-derived cells into osteoblasts (Frenkel et al. *Endocrinology*
20 (in press)). Such cells can thus be used to confirm that a specific fragment of a promoter is inactive in some daughter cells and active in other daughter cells. Progenitor cells capable of differentiating into various types of cells upon being submitted to the proper conditions have also been established as cell lines and are available from biological depositories such as the American Type Culture Collection.

25 Other assays for determining the activity of regulatory elements, such as bone specific promoters include those in which cultures of primary bone cells, such as fetal calvarial bone cells, are transfected with the promoter of interest linked to a reporter gene. Upon culture of these cells under appropriate conditions and which are known in the art, the cells will proliferate and eventually develop into nodules, which then become
30 mineralized. Thus, the cells undergo progressive differentiation to the stage of osteocytes. Such systems are described for example in Owen et al. (1990) *J. Cellular Physiology* 143: 420-430. These systems allow monitoring of the activity of the regulatory element during development of the osteoblast cells.

Alternatively, the transcriptional activity of a regulatory element can be assayed
35 by preparing transgenic mice containing the specific element as the transgene. Transgenic mice can be prepared according to methods known in the art and which are described in Example 1 and also, e.g., in Leder et al. (U.S. Patent No. 4,736,866).

Expression of the exogenous gene in the expected tissues of the transgenic mouse will indicate that the regulatory element is functionally active and can be used in the methods of the invention.

5 **Exogenous genes**

 The invention provides methods for expressing an exogenous gene of interest in a specific tissue or cell population by transfecting pluripotent stem cells with the exogenous gene under the control of appropriate regulatory sequences. Accordingly, the invention provides methods for expressing a specific gene of interest in a specific tissue
10 or cell population. Such methods can be used for various gene therapy protocols, in which it is desirable to obtain tissue-specific expression of a gene. In a preferred embodiment, the invention provides methods for expressing a specific gene in bone tissue.

 The term exogenous gene is intended to include any gene or fragment thereof, or
15 modification thereof which is introduced into a cell. An exogenous gene of the invention can encode a protein or a peptide. An exogenous gene of the invention can also be a nucleic acid that is transcribed into RNA, but does not encode a peptide. For example, an exogenous gene can be a nucleic acid which, upon transcription into an RNA molecule is an "antisense" strand of another nucleic acid in or out of the cell, such
20 that upon expression of the exogenous gene and synthesis of antisense molecules, a function in the cell is modulated. In another embodiment of the invention, the antisense nucleic acid inhibits or reduces expression of another nucleic acid, such as an endogenous nucleic acid.

 In another embodiment, the exogenous gene encodes a therapeutic protein useful
25 for treating a disease or condition. The exogenous gene can encode a secreted protein, a membrane bound protein, or an intracellular protein. Preferred exogenous genes encode a therapeutic protein. A therapeutic protein can be a steroid hormone, a steroid hormone receptor, a growth factor, a cytokine, a morphogenic protein, a polypeptide hormone, a polypeptide chemotherapeutic agent, a signal transduction factor and an intermediate.
30 Preferred morphogenic proteins include bone morphogenic proteins (BMPs). Other preferred exogenous genes include multidrug resistance genes and genes encoding calcitonin or collagen components. Expression of multidrug resistance genes, e.g., MDR1, in bone cells should provide host resistance to a variety of chemotherapeutic drugs.

35 Introduction into an individual of bone marrow transfected according to the method of the invention can be used to obtain bone specific expression of one or more proteins which modulate bone growth, e.g., stimulate bone growth. Accordingly, bone

growth can be stimulated according to the method of the invention, wherein the exogenous gene encodes a growth factor.

In another embodiment, the exogenous gene encodes a suicide gene. A suicide gene is intended to include genes which induce cell death when the cell is exposed to a specific condition. A suicide gene can also be a gene which causes death of cells located in close proximity to the cells expressing the suicide gene. Examples of suicide genes include genes encoding a protein involved in apoptosis. Such genes include Fas. Apoptosis in Fas-bearing cells is induced upon binding of a Fas ligand with the Fas receptor on the Fas-bearing cells. (Nagata, S. and Golstein, P. (1995) *Science* 267: 1449). Other apoptotic genes include genes from the Bcl-2 family, such Bcl-Xs (Boise, L., et al. (1993) *Cell* 74, 597-608; Martin, S. and Green, D. (1995) *Critical Reviews in Oncology/Hematology* 18: 137-153; and Savill, J. (1994) *European Journal of Clinical Investigation* 24: 715-723). Another suicide gene is the thymidine kinase (TK) gene. A cell expressing TK will be killed upon treatment with gancyclovir or acyclovir. For example, Ko et al. ((1996) *Cancer Res.* 56:4614) have shown that *in vivo* TK gene expression in mice in murine osteosarcoma cells in the presence of acyclovir is effective in inhibiting murine osteosarcoma growth. In another embodiment of the invention, the suicide gene is a toxin, such as ricin.

Cells located in sufficient close proximity to cells expressing the suicide gene can be killed by the "bystander effect". In fact, it has been demonstrated recently that tumor elimination or tumor regression can occur without expression of the suicide gene in every tumor cell *in vivo* (Tanaka et al. (1996) *Cancer Res.* 46:1341; Chen et al. (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92:2577; Smythe et al. (1995) *Ann. Surg.* 222:78). Accordingly, targeting of a bone tumor with bone cells expressing a suicide gene, can result in tumor elimination or regression. Similarly, this method of tumor cell killing can be used for treating other cancers, e.g., hepatoma, glioma, melanoma and gastric carcinomas. Specific promoters that can be used for treating these cancers are described, e.g., in Kuriyama et al. (1991) *Cell Struct. Funct.* 16:503, Shimizu (1994) *Nippon Rinsho* 52:3053, Vile et al. (1994) *Cancer Res.* 54:6228, and Tanaka, *supra*).

In a specific embodiment of the invention, the exogenous gene comprises a nucleotide sequence containing one or more open reading frames, i.e., sequences that code for peptides, such that upon transfection into the bone marrow cells according to the method of the invention, at least one protein is synthesized in the target cell deriving from a bone marrow cell. The gene encoding at least one protein can be any gene, such as a gene encoding a cytokine. The gene can code for one peptide or the gene can encode several peptides.

- 17 -

In another embodiment of the invention, the exogenous gene is a nucleotide sequence which is expressed as one or more functional RNA molecules (e.g., an antisense RNA molecule). In a preferred embodiment of the invention, the functional RNA molecule inhibits, expression of one or more endogenous genes in the specific target cell. Thus, the method of the invention is useful for decreasing expression of a selected gene in cells of a specific tissue. For example, bone marrow cells can be transfected with a nucleic acid molecule comprising a gene encoding antisense RNA, such that translation of an endogenous RNA is reduced. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid, e.g., complementary to an mRNA sequence encoding a protein, constructed according to the rules of Watson and Crick base pairing. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense sequence complementary to a sequence of an mRNA can be complementary to a sequence found in the coding region of the mRNA or can be complementary to a 5' or 3' untranslated region of the mRNA. Preferably, an antisense nucleic acid is complementary to a region preceding or spanning the initiation codon or in the 3' untranslated region of an mRNA. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. et al. (1986) *Trends in Genetics*, Vol. 1(1).

In another embodiment of the invention, expression of an endogenous gene in a specific tissue or cell population is reduced by introducing into the bone marrow cells a nucleic acid encoding a ribozyme according to the method of the invention. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. A ribozyme having specificity for a nucleic acid of interest can be designed based upon the nucleotide sequence of the nucleic acid. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the base sequence of the active site is complementary to the base sequence to be cleaved in an mRNA of interest. See for example Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742.

The "nucleic acid molecule" to be transfected into the bone marrow cells comprising the gene can be a DNA molecule or an RNA molecule. The nucleic acid molecule can be a portion of a natural nucleic acid molecule, or alternatively, it can be made synthetically.

Methods for obtaining pluripotent stem cells from a subject

The invention pertains to methods for obtaining tissue-specific expression of an exogenous gene, e.g., a therapeutic gene, in which pluripotent stem cells are transfected

with the exogenous gene under the control of at least one tissue-specific regulatory element. Sources for pluripotent stem cells for use in the present invention include bone marrow, cord blood and mobilized peripheral blood (MPB). In preferred embodiments, the pluripotent stem cell is isolated from a mammal, e.g., a primate, e.g., a human. The invention also includes the use of stem cells from transgenic non-human mammals. In one embodiment, the pluripotent stem cells are obtained from the subject into which the stem cells are to be transplanted after *in vitro* culturing and transduction of the exogenous gene.

The source of cells for the present invention can be, in addition to humans, non-human mammals. A variety of protocols are known in the art for isolating pluripotent stem cells from non-human animals. See, for example, the Wheeler U.S. Patent 5,523,226 entitled "Transgenic swine compositions and methods" and the Emery et al. PCT publication WO 95/13363 entitled "Hematopoietic Stem Cells From Swine Cord Blood And Uses Thereof". The preferred non-human animals include vertebrates such as rodents, non-human primates, sheep, dog, cow and pigs. The term "non-human mammal" refers to all members of the class *Mammalia* except humans.

As used herein, a "transgenic animal" is any animal, preferably a non-human mammal, bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. Moreover, "transgenic animal" also includes those recombinant animals in which gene disruption of one or more genes is caused by human intervention, including both recombination and antisense techniques.

Where the intended use of the resulting transduced cell is for transplantation in human patients, the cells derived from transgenic animals can be used as a source for "humanized" hematopoietic cells, e.g., for xenogenic grafting into human subjects. For example, as described by the Sachs et al. PCT publication WO 96/06165 entitled "Genetically Engineered Swine Cells", the art provides for transplantation of swine donor cells which have been engineered to increase desirable interactions between the donor cells and molecules and cells of a recipient, e.g., to promote the engraftment or function of the donor stem cells in the recipient environment. To illustrate, the cells can be engineered to express a human adhesion molecule, e.g., an adhesion molecule

involved in engraftment and/or maintenance of transplanted cells. Examples of human adhesion molecules include VLA-4, c-kit, LFA-1, CD11a, Mac-1, CR3, CD11b, p150, p95, CD11c, CD49a, LPAM-1, CD49d, CD44, CD38, and CD34. The transgenic cells can also be engineered to minimize unwanted interactions between the donor cells and molecules and cells of the recipient which, e.g., promote the rejection of donor graft cells or which inhibit the function of the donor graft cells. For example, the donors cells can be derived from a transgenic animal expressing one or more human MHC polypeptides.

Procedures for obtaining bone marrow which contain pluripotent stem cells are known by those skilled in the art and are described in a variety of medical textbooks. For example, bone marrow cells can be obtained from a source of bone marrow, including but not limited to, ilium (e.g. from the hip bone via the iliac crest), tibia, femora, spine, or other bone cavities. Other sources of stem cells include, but are not limited to, embryonic yolk sac, fetal liver, and fetal spleen.

For isolation of bone marrow, an appropriate solution can be used to flush the bone, e.g., a salt solution supplemented with fetal calf serum (FCS) or other naturally occurring factors, in conjunction with an acceptable buffer at low concentration, generally from about 5-25 mM. Convenient buffers include HEPES, phosphate buffers and lactate buffers. Otherwise bone marrow can be aspirated from the bone in accordance with conventional techniques. The bone marrow harvests are preferably maintained in anticoagulation media, such as media containing about 10,000 units preservative-free heparin and about 50 cc anticoagulant (ACD) per about 100 cc tissue culture media. About 450 cc of bone marrow harvest is preferably added to about 50 cc of this media to which another about 50 cc of ACD is added.

Methods for mobilizing stem cells into the peripheral blood are known in the art and generally involve treatment with chemotherapeutic drugs, cytokines (e.g. GM-CSF, G-CSF or IL3), or combinations thereof. Typically, apheresis for total white cells begins when the total white cell count reaches 500-200 cells/l and the platelet count reaches 50,000/l.

Fetal or neonatal blood are also sources for the stem and progenitor cells used in the present invention. Fetal blood can be obtained by any method known in the art. For example, fetal blood can be taken from the fetal circulation at the placental root with the use of a needle guided by ultrasound (Daffos et al., (1985) *Am. J. Obstet Gynecol* 153:655-660; Daffos et al., (1983) *Am. J. Obstet. Gynecol.* 146:985), by placentocentesis (Valenti, C., (1973) *Am. J. Obstet. Gynecol.* 115:851; Cao et al., (1982) *J. Med. Genet.* 19:81), by fetoscopy (Rodeck, C.H., (1984) in Prenatal Diagnosis, Rodeck, C.H. and Nicolaides, K.H., eds., Royal College of Obstetricians and Gynaecologists, London), etc.

In one embodiment of the invention, neonatal pluripotent stem and progenitor cells can be obtained from umbilical cord blood and/or placental blood. The use of cord or placental blood as a source of stem cells provides numerous advantages. Cord blood can be obtained easily and without trauma to the donor. In contrast, at present, the collection of bone marrow cells is a traumatic experience which is costly in terms of time and money spent for hospitalization. Cord blood cells can be used for autologous transplantation, when and if needed, and the usual hematological and immunological problems associated with the use of allogeneic cells, matched only partially at the major histocompatibility complex or matched fully at the major, but only partially at the minor complexes, are alleviated.

Cell collections should be made under sterile conditions. Immediately upon collection, the neonatal or fetal blood should be mixed with an anticoagulant. Such an anticoagulant can be any known in the art, including but not limited to CPD (citrate-phosphate-dextrose), ACD (acid citrate-dextrose), Alsever's solution, De Gowin's Solution, Edglugate-Mg, Rous-Turner Solution, other glucose mixtures, heparin, ethyl biscoumacetate, etc. (See Hurn, B.A.L., 1968, Storage of Blood, Academic Press, New York, pp. 26-160).

In those embodiments where a purification step is not carried out, tissue-specific expression of an exogenous gene is obtained at least in part by the tissue-specific regulatory elements which drive the expression of those exogenous genes. Therefore, it is not necessary to purify the population of bone marrow cells either before or after transfection with the exogenous gene and the regulatory elements.

Alternatively, a population of purified pluripotent stem cells, e.g., osteoprogenitor cells, can be used. As used herein, the terms "purified" or "enriched" refer to a population of stem cells that is at least about 60%, preferably at least about 70%, more preferably at least about 80%, and most preferably at least about 90% pure, with respect to a total cell population.

Various techniques can be employed to separate the cells by initially removing lineage committed cells. Monoclonal antibodies are particularly useful for identifying markers associated with particular cell lineages and/or stages of differentiation. Antibodies which find use include antibodies to lineage specific markers which allow for removal of most, if not all, mature cells, while being absent on stem cells. The antibodies can be attached to a solid support to allow for crude separation. The separation techniques employed should maximize the retention of viability of the fraction to be collected. Various techniques of different efficacy can be employed to obtain "relatively crude" separations. Such separations are where up to 10%, usually not more than about 5%, preferably not more than about 1%, of the total cells present not

- 21 -

having the marker can remain with the cell population to be retained. The particular technique employed will depend upon efficiency of separation, associated cytotoxicity, ease and speed of performance, and necessity for sophisticated equipment and/or technical skill.

5 The use of separation techniques include, but are not limited to, those based on differences in physical (density gradient centrifugation and counter-flow centrifugal elutriation), cell surface (lectin and antibody affinity), and vital staining properties (mitochondria-binding dye rho123 and DNA-binding dye Hoechst 33342). Procedures for separation can include, but are not limited to, magnetic separation, using antibody-coated magnetic beads, affinity chromatography, cytotoxic agents joined to a
10 monoclonal antibody or used in conjunction with a monoclonal antibody, including, but not limited to, complement and cytotoxins, and "panning" with antibody attached to a solid matrix, e.g., plate, elutriation or any other convenient technique. Techniques providing accurate separation include, but are not limited to, FACS, which can have
15 varying degrees of sophistication, e.g., a plurality of color channels, low angle and obtuse light scattering detecting channels, impedance channels, etc. Concomitantly or subsequent to a gross separation, which provides for positive selection, a negative selection can be carried out, where antibodies to lineage-specific markers present on dedicated cells are employed.

20 While it is believed that the particular order of separation is not critical to this invention, the order indicated is preferred. Preferably, cells are initially separated by a coarse separation, followed by a fine separation, with positive selection of a marker associated with stem cells and negative selection for markers associated with lineage committed cells.

25

Methods for transfecting a population of pluripotent stem cells with an exogenous gene

 The population of pluripotent stem cells, e.g., obtained from a subject as described above, can be transfected with a nucleic acid comprising an exogenous gene under the control of at least one regulatory element using a variety of methods, some of
30 which are set forth below.

 The terms "transfection" or "transfected with" refers to the introduction of exogenous nucleic acid into a mammalian cell and encompass a variety of techniques useful for introduction of nucleic acids into mammalian cells including electroporation, calcium-phosphate co-precipitation, DEAE-dextran treatment, liposome-mediated gene
35 transfer, microinjection and infection with viral vectors. Suitable methods for transfecting mammalian cells can be found in Sambrook et al. (Molecular Cloning: A

Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory press (1989)) and other laboratory textbooks.

For transfection of an exogenous gene and regulatory sequences into pluripotent stem cells, it is preferable that these nucleic acids be contained in a plasmid or vector containing sequences or elements well known in the art for preparing the nucleic acid prior to transfection. Such sequences include those that enable the nucleic acid to be replicated, such as a bacterial origin of replication. Suitable plasmid expression vectors include CDM8 (Seed, B., *Nature* 329, 840 (1987)) and pMT2PC (Kaufman, et al., *EMBO J.* 6, 187-195 (1987)). It may be desirable to select for the bone marrow cells which have incorporated the nucleic acid after the transfection. This can be performed, e.g., by transfecting a nucleic acid encoding a selectable marker into the bone marrow cells along with the nucleic acid(s) of interest. Preferred selectable markers include those which confer resistance to drugs such as G418, hygromycin and methotrexate. Selectable markers may be introduced on the same plasmid as the gene(s) of interest or may be introduced on a separate plasmid. Following selection of transfected cells using the appropriate selectable marker(s), expression of the exogenous gene can be confirmed by various methods including immunofluorescent staining of the cells and measure of a biological activity of the protein encoded by the exogenous gene.

In a preferred embodiment of the invention, the exogenous gene linked to a regulatory element is introduced into pluripotent stem cells using a viral vector. Such viral vectors include, for example, recombinant retroviruses, adenovirus, adeno-associated virus, and herpes simplex virus-1. Viral vectors are generally understood to be the recombinant gene delivery system of choice for the transfer of exogenous genes *in vivo*, particularly into humans. Alternatively, they can be used for introducing exogenous genes *ex vivo* into cells. These vectors provide efficient delivery of genes into cells, and the transferred nucleic acids are stably integrated into the chromosomal DNA of the host cell. A major prerequisite for the use of retroviruses is to ensure the safety of their use, particularly with regard to the possibility of the spread of wild-type virus in the cell population. The development of specialized cell lines (termed "packaging cells") which produce only replication-defective retroviruses has increased the utility of retroviruses for gene therapy, and defective retroviruses are well characterized for use in gene transfer for gene therapy purposes (for a review see Miller, A.D. (1990) *Blood* 76:271). Thus, recombinant retrovirus can be constructed in which part of the retroviral coding sequence (*gag*, *pol*, *env*) is replaced by a gene of interest rendering the retrovirus replication defective. The replication defective retrovirus is then packaged into virions which can be used to infect a target cell through the use of a helper virus by standard techniques. Protocols for producing recombinant retroviruses and for

infecting cells *in vitro* or *in vivo* with such viruses can be found in Current Protocols in Molecular Biology, Ausubel, F.M. et al. (eds.) Greene Publishing Associates, (1989), Sections 9.10-9.14 and other standard laboratory manuals. Examples of suitable retroviruses include pLJ, pZIP, pWE and pEM which are well known to those skilled in the art. Examples of suitable packaging virus lines for preparing both ecotropic and amphotropic retroviral systems include ψ Crip, ψ Cre, ψ 2 and ψ Am.

Furthermore, it has been shown that it is possible to limit the infection spectrum of retroviruses and consequently of retroviral-based vectors, by modifying the viral packaging proteins on the surface of the viral particle (see, for example PCT publications WO93/25234 and WO94/06920). For instance, strategies for the modification of the infection spectrum of retroviral vectors include: coupling antibodies specific for cell surface antigens to the viral *env* protein (Roux et al. (1989) *PNAS* 86:9079-9083; Julian et al. (1992) *J. Gen Virol* 73:3251-3255; and Goud et al. (1983) *Virology* 163:251-254); or coupling cell surface receptor ligands to the viral *env* proteins (Neda et al. (1991) *J Biol Chem* 266:14143-14146). Coupling can be in the form of the chemical cross-linking with a protein or other variety (e.g. lactose to convert the *env* protein to an asialoglycoprotein), as well as by generating fusion proteins (e.g. single-chain antibody/*env* fusion proteins). Thus, in a specific embodiment of the invention, viral particles containing a nucleic acid molecule containing a gene of interest, i.e., an exogenous gene, such as a therapeutic gene, operably linked to appropriate regulatory elements, are modified for example according to the methods described above, such that they can specifically target subsets of the bone marrow cells. For example, the viral particle can be coated with antibodies to surface molecule that are specific to certain types of progenitor cells.

Another viral gene delivery system useful in the present invention utilizes adenovirus-derived vectors. The genome of an adenovirus can be manipulated such that it encodes and expresses a gene product of interest but is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. See for example Berkner et al. (1988) *BioTechniques* 6:616; Rosenfeld et al. (1991) *Science* 252:431-434; and Rosenfeld et al. (1992) *Cell* 68:143-155. Suitable adenoviral vectors derived from the adenovirus strain Ad type 5 dl324 or other strains of adenovirus (e.g., Ad2, Ad3, Ad7 etc.) are well known to those skilled in the art. Recombinant adenoviruses can be advantageous in certain circumstances in that they are not capable of infecting nondividing cells. Furthermore, the virus particle is relatively stable and amenable to purification and concentration, and as above, can be modified so as to affect the spectrum of infectivity. Additionally, introduced adenoviral DNA (and foreign DNA contained therein) is not integrated into the genome of a host cell but remains episomal,

thereby avoiding potential problems that can occur as a result of insertional mutagenesis in situations where introduced DNA becomes integrated into the host genome (e.g., retroviral DNA). Moreover, the carrying capacity of the adenoviral genome for foreign DNA is large (up to 8 kilobases) relative to other gene delivery vectors (Berkner et al. cited *supra*; Haj-Ahmand and Graham (1986) *J. Virol.* 57:267). Most replication-defective adenoviral vectors currently in use and therefore favored by the present invention are deleted for all or parts of the viral E1 and E3 genes but retain as much as 80 % of the adenoviral genetic material (see, e.g., Jones et al. (1979) *Cell* 16:683; Berkner et al., *supra*; and Graham et al. in Methods in Molecular Biology, E.J. Murray, Ed. (Humana, Clifton, NJ, 1991) vol. 7. pp. 109-127). Expression of the gene of interest comprised in the nucleic acid molecule can be under control of, for example, the E1A promoter, the major late promoter (MLP) and associated leader sequences, the E3 promoter, or exogenously added promoter sequences.

Yet another viral vector system useful for delivery of a nucleic acid molecule comprising a gene of interest is the adeno-associated virus (AAV). Adeno-associated virus is a naturally occurring defective virus that requires another virus, such as an adenovirus or a herpes virus, as a helper virus for efficient replication and a productive life cycle. (For a review see Muzyczka et al. *Curr. Topics in Micro. and Immunol.* (1992) 158:97-129). Adeno-associated viruses exhibit a high frequency of stable integration (see for example Flotte et al. (1992) *Am. J. Respir. Cell. Mol. Biol.* 7:349-356; Samulski et al. (1989) *J. Virol.* 63:3822-3828; and McLaughlin et al. (1989) *J. Virol.* 62:1963-1973). Vectors containing as few as 300 base pairs of AAV can be packaged and can integrate. Space for exogenous DNA is limited to about 4.5 kb. An AAV vector such as that described in Tratschin et al. (1985) *Mol. Cell. Biol.* 5:3251-3260 can be used to introduce DNA into bone marrow cells. A variety of nucleic acids have been introduced into different cell types using AAV vectors (see for example Hermonat et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6466-6470; Tratschin et al. (1985) *Mol. Cell. Biol.* 4:2072-2081; Wondisford et al. (1988) *Mol. Endocrinol.* 2:32-39; Tratschin et al. (1984) *J. Virol.* 51:611-619; and Flotte et al. (1993) *J. Biol. Chem.* 268:3781-3790). Other viral vector systems that may have application in gene therapy have been derived from herpes virus, vaccinia virus, and several RNA viruses.

Transfection of pluripotent stem cells using viral vectors can be performed as described, e.g., in Eglitis et al. (1985) *Science* 230:1395, Kohn et al. (1987) *Blood Cells* 13:285, and Kantoff et al. (1987) *J. Exp. Med.* 166:219.

In another embodiment of the invention, the nucleic acid molecule comprising a gene of interest is introduced into pluripotent stem cells by non-viral-mediated methods

of transfection well known in the art. These methods include electroporation, calcium phosphate precipitation, and DEAE dextran transfection.

In yet another embodiment, the nucleic acid molecule comprising an exogenous gene operably linked to appropriate regulatory elements is carried by and delivered into
5 bone marrow cells by a cell-delivery vehicle. Such vehicles include, for example, cationic liposomes (LipofectinTM) or derivatized (e.g. antibody conjugated) polylysine conjugates, gramicidin S, artificial viral envelopes. These vehicles can deliver a nucleic acid that is incorporated into a plasmid, vector, or viral DNA.

In another embodiment of the invention, the nucleic acid molecule comprising a
10 gene of interest is delivered into a specific cell in the form of a soluble molecular complex. The complex contains the nucleic acid releasably bound to a carrier comprised of a nucleic acid binding agent and a cell-specific binding agent which binds to a surface molecule of the specific T cell and is of a size that can be subsequently internalized by the cell. Such complexes are described, e.g., in U.S. Patent Serial No. 5,166,320.

15 In another embodiment of the invention the nucleic acid is introduced into pluripotent stem cells by particle bombardment, as described in Yang, N.-S. and Sun, W.H. (1995) *Nature Medicine* 1, 481.

Following transfection of the pluripotent stem cells with the exogenous gene and regulatory elements together or not with a selection marker and selection, if appropriate,
20 and demonstration that the exogenous gene is expressed in at least some cells, the pluripotent stem cells can be administered to a subject.

The population of pluripotent stem cells either prior to, or after transfection, can be frozen and stored for prolonged time. Alternatively, the pluripotent stem cells are obtained from a subject, transfected *ex vivo*, and administered to a subject, e.g., the
25 subject from which the cells were obtained originally.

Methods for administering transfected bone marrow cells to a subject

The transfected pluripotent stem cells can be administered to a subject either locally or systemically. As used herein, the term "subject" is intended to include human
30 and non-human animals. The term "non-human animals" of the invention includes all vertebrates, e.g., mammals and non-mammals, such as non-human primates, sheep, dog, cow, chickens, amphibians, reptiles, etc. According to the method of the invention, osteoprogenitor cells carrying the exogenous gene engraft and differentiate into mature osteoblasts, thereby expressing the therapeutic gene. Other cells are also transfected and
35 engraft, but do not express the transfected gene because they do not support activity of the osteocalcin promoter.

- 26 -

Bone marrow transplant experiments which contain pluripotent stem cells demonstrate that the stem cells home to marrow cavities of either irradiated or non-myeloablated recipient animals without a requirement for preparing an engraftment space (Stewart et al. (1993) *Blood* 81:2566). Accordingly, it is not necessary to destroy
5 bone marrow cells of a subject prior to administering the transfected bone marrow of the invention to the subject. Transplantation without myeloablation is desirable in certain conditions, e.g, in cases of genetic diseases or HIV infection.

Methods for administering bone marrow transplants to a subject are known in the art and are described in medical textbooks, e.g., Whedon, M.B. (1991) Whedon, M.B.
10 "Bone Marrow Transplantation: Principles, Practice, and Nursing Insights", Boston: Jones and Bartlett Publishers. In certain embodiments, bone marrow cells from a healthy patient can be removed, preserved, and then replicated and reinfused should the patient develop an illness which either destroys the bone marrow directly or whose treatment adversely affects the marrow. If the patient is receiving his or her own cells,
15 this is called an autologous transplant; such a transplant has little likelihood of rejection. Autologous transplants eliminate a major cause of bone marrow transplant rejection, that is, the graft vs. host reaction. If the marrow contains malignant or diseased cells, small samples of it can be more effectively purged and the stem cells expanded using culture system known in the art. As understood in the art, selective methods for purging
20 malignant or diseased cells work best in small volumes of bone marrow cells.

Pharmaceutical Preparation of Cells

In one aspect, the methods of the present invention provide a population of stem cells transfected *ex vivo* with an exogenous gene. The transfected stem cells can be
25 administered to a subject. Exemplary methods of administering the stem cells to subjects, particularly human subjects, include injection or transplantation of the cells into target sites in the subjects. The cells produced by the methods of the invention can be inserted into a delivery device which facilitates introduction by, injection or transplantation, of the cells into the subjects. Such delivery devices include tubes, e.g.,
30 catheters, for injecting cells and fluids into the body of a recipient subject. In a preferred embodiment, the tubes additionally have a needle, e.g., a syringe, through which the cells of the invention can be introduced into the subject at a desired location. The stem cells can be inserted into such a delivery device, e.g., a syringe, in different forms. For example, the cells can be suspended in a solution or embedded in a support matrix when
35 contained in such a delivery device.

As used herein, the term "solution" includes a pharmaceutically acceptable carrier or diluent in which the cells of the invention remain viable. Pharmaceutically

- 27 -

acceptable carriers and diluents include saline, aqueous buffer solutions, solvents and/or dispersion media. The use of such carriers and diluents is known in the art. The solution is preferably sterile and fluid to the extent that easy syringability exists.

Preferably, the solution is stable under the conditions of manufacture and storage and preserved against the contaminating action of microorganisms such as bacteria and fungi through the use of, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. Solutions of the invention can be prepared by incorporating stem cells as described herein in a pharmaceutically acceptable carrier or diluent and, as required, other ingredients enumerated above, followed by filtered sterilization.

Support matrices in which the stem cells can be incorporated or embedded include matrices which are recipient-compatible and which degrade into products which are not harmful to the recipient. Natural and/or synthetic biodegradable matrices are examples of such matrices. Natural biodegradable matrices include plasma clots, e.g., derived from a mammal, and collagen matrices. Synthetic biodegradable matrices include synthetic polymers such as polyanhydrides, polyorthoesters, and polylactic acid. Other examples of synthetic polymers and methods of incorporating or embedding cells into these matrices are known in the art. See e.g., U.S. Patent No. 4,298,002 and U.S. Patent No. 5,308,701. These matrices provide support and protection for the fragile stem cells *in vivo* and are, therefore, the preferred form in which the stem cells are introduced into the recipient subjects.

Uses for the invention

The invention provides methods for expressing an exogenous gene, such as a therapeutic gene, in selected cell populations or tissues. Accordingly, the invention can be used for treating any condition in which production of a specific protein in a specific tissue or cell type is desirable. Thus, the invention provides a gene therapy tool that provides tissue-specific expression of a gene without requiring extensive purification of a specific cell population.

In one embodiment, the invention is a new process for treating primary and secondary bone disorders with therapeutic genes targeted to be expressed predominantly in skeletal tissue.

In another embodiment, stem cells, e.g., bone marrow cells, are transfected *ex-vivo* with an exogenous gene, e.g., a therapeutic gene, under the control of a bone-specific osteocalcin gene regulatory element as described herein. For example, transfected marrow-derived cells can then be transplanted into a host, e.g., using a non-myeloablation protocol. Consequently, expression of the exogenous gene will be confined to functional osteoblasts that develop from marrow-derived mesenchymal stem

cells, and high local concentrations of the exogenous gene product will therefore be restricted to skeletal tissue.

In yet another embodiment, the invention provides methods for treating metabolic bone diseases, skeletal disorders or malignancies. Such skeletal disorders
5 include osteoporosis (including post-menopausal osteoporosis), osteopenia (including drug-induced osteopenia), osteosarcoma, metastasis, and osteomalaciae. The invention also provides methods for treating osteosarcomas and other bone neoplasiae. The invention further provides methods for treating non-osseous tumors that metastasize to bone (e.g., breast cancer and prostate cancer). According to a preferred method of the
10 invention, osteosarcomas and neoplasiae can be treated by selectively expressing a suicide gene in the malignant cells. The invention also provides methods for treating traumatic and iatrogenic bone lesions.

Other diseases which can be treated according to the method of the invention include drug-induced osteomalacia, fibrous dysplasia, osteogenesis imperfecta, and
15 Paget's disease of bone.

Also within the scope of the invention are diseases of collagen, e.g., osteogenesis imperfecta and chondrodysplasias. The method of the invention can be used to provide corrected version of collagen genes in cases where a disease arises from the existence of a mutated form of a collagen molecule.

20 The invention also provides methods for treating disorders of tissues other than bone tissues, utilizing tissue-specific regulatory elements, such as tissue-specific promoters, e.g., collagen type II for the treatment of cartilage disorders.

In one embodiment in which an exogenous gene is under the control of a heat sensitive element, such as a heat shock element, it is possible to obtain expression of the
25 exogenous gene in a subject with a fever.

At least one advantage of the methods of the invention for treating bone diseases which include, but are not restricted to osteoporosis, immunosuppression drug-induced osteopenia, osteosarcoma, and a series of non-osseous primary tumors (e.g., breast cancer, prostate cancer) which metastasize to bone is that the methods of the invention
30 minimize the extent of invasive surgical procedures or toxic drug treatment involved. Thus, the methods of the invention reduce the level of risk to the patient as compared to classical methods of treating bone diseases.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patent, and published patent
35 applications cited throughout this application are hereby incorporated by reference.

Example 1: Activity of the rat osteocalcin promoter in transgenic mice

This example demonstrates that high tissue-specific expression can be obtained with the osteocalcin promoter.

Transgenic mice carrying the chloramphenicol acetyl transferase (CAT) reporter gene under the control of various portions of the rat osteocalcin (OC) 5' flanking sequence. Portions of about 0.35 kb, 0.72 kb, 1.1 kb, and 1.7 kb of 5' flanking sequence were tested. More specifically, these portions consisted of -348/+24 (0.35 kb construct), -724/+23 (0.72 kb construct), -1097/+23 (1.1 kb construct), and -1727/+23 (1.7 kb construct) fragments of the rat osteocalcin gene shown in Figure 3 (SEQ ID NO: 18).

Each of these four constructs exhibited CAT expression in bone in at least one transgenic line. Thus, as little as 0.35 kb of osteocalcin promoter can direct bone expression *in vivo*. CAT activity in bone was highest in a transgenic line carrying a 1.7 kb promoter construct (54% conversion/hr/mg in the femur) as compared to mice carrying 1.1, 0.72, and 0.35 kb promoter constructs.

The mouse line harboring the 1.7 kb promoter construct was expanded and offspring ranging from 5 weeks to 11 months were assayed for CAT activity in muscle, liver, spleen, kidney, lung, brain, femur, calvaria, lumbar vertebrae and bone marrow. CAT activity was detected only in the skeletal sites with the exception of brain (~1% of bone level). The amount of CAT in one milligram tissue extract, derived from the femur or the lumbar vertebrae acetylated about 2 nmol of chloramphenicol per hour at 37 °C (with initial concentrations of 0.1 mM and 1 mM of chloramphenicol and acetyl CoA, respectively).

An age-related decline in osteocalcin promoter activity was observed by three months, and notably, females exhibited a higher level of CAT activity compared to males.

Activity in the bone marrow was negligible at all ages, reflecting the absence of mature osteoblasts in this compartment. However, when bone marrow cells derived from 3-month old transgenic mice were cultivated for 2 weeks in the presence of 20% fetal calf serum, high CAT activity developed in mineralizing cell cultures, reflecting the differentiation of bone marrow mesenchymal stem cells to mature osteoblasts. Both this differentiation process and OC 5' flanking-driven CAT activity were inhibited by dexamethasone (10^{-7} M), reflecting fidelity of OC 5' flanking sequence activity in relation to osteoblast differentiation.

The above-described *in vitro* differentiation assay was performed as follows. Mice were sacrificed by cervical dislocation and the femurs were aseptically removed. The epiphyses and growth plates were dissected and the marrow flushed in cold culture medium (α MEM) supplemented with heat inactivated 20% fetal bovine serum (Atlanta

- 30 -

Biologicals, Norcross, GA) and 10^{-8} M menadione sodium bisulfate (Sigma, St. Louis, MI). Cells were centrifuged for 10 minutes at 1000g, resuspended in fresh culture medium, filtered through a 100-mesh steel screen and seeded onto 35 mm 6 well plates at 10^7 cells/well. Ascorbic acid (50 μ g/ml) and β -glycerophosphate (10 mM) were
5 added 24 hours later and were present throughout the culture period. Medium was changed on day 4 and every 48 hours thereafter.

Thus, this example shows that a fragment of 1.7 kb from the 5' flanking sequence of the osteocalcin gene provides high transcriptional activity and predominant expression in skeletal tissue. This example further demonstrates the utility of this
10 transgenic model for selecting portions of the osteocalcin 5' flanking sequence that results in high levels of tissue-specific transcription.

Example 2: Transplantation of bone marrow cells into mice

This example demonstrates that bone marrow cells can be obtained from a
15 mouse, transfected *ex vivo* and reintroduced into mice.

For 6-8 week old mice, five consecutive doses of bone marrow cells are administered daily via the tail vein. Each dose contains 40 million cells in 0.5-1.0 ml phosphate-buffered saline as described in Stewart et al. (1993) *Blood* 81:2566.

Example 3: Expression of osteocalcin promoter-driven genes in expanded populations of transplanted bone marrow cells

This example demonstrates that expanded populations of whole bone marrow-derived cells engraft and express an osteocalcin promoter-driven gene in bone tissues of a recipient animal.

25 Whole bone marrow cells were obtained from a 6 week old transgenic mouse harboring a 1.7 kb rat osteocalcin promoter-CAT reporter construct, which is expressed in the mouse primarily in osseous tissues (see Example 1). The cells were expanded in culture in α -MEM supplemented with 20% heat-inactivated fetal calf serum. The cells were trypsinized at any time from day 8 to 13, inclusive, and 4×10^6 cells were
30 introduced into a 7 week old recipient mouse by tail vein injection. Analysis of the recipient mouse eight weeks after injection revealed detectable CAT activity in the calvarium and the femur reflecting bone tissue-specific activity of the osteocalcin promoter in the transplanted cells. Analysis of the transplanted bone marrow cells by genomic DNA polymerase chain reaction (PCR) demonstrated that the transplanted cells
35 engrafted into skeletal and non-skeletal tissues. However, gene expression was confined to osteoblasts and osteocytes as determined by histochemical detection of the CAT reporter.

Example 4: Tissue-Specific Expression of osteocalcin promoter-driven genes in transplanted bone marrow cells

Genomic DNA polymerase chain reaction (PCR) was used to determine the location of engrafted donor cells carrying the osteocalcin promoter fused to a CAT reporter in recipient mice. Mice were transplanted as described Example 2. To selectively detect donor cells, PCR primers were designed which recognize DNA sequences in both the osteocalcin promoter (-551 bp) and the CAT reporter (+596 bp). The results of PCR reaction products demonstrated that the donor cells engrafted and remained in both osseous and non-osseous tissues for up to 10 months post transplantation. In contrast to the widespread location of transplanted cells, the osteocalcin promoter was operative predominantly in bone cells. Osteocalcin promoter activity was assessed by detecting expression of the CAT reporter gene product (CAT protein) using standard techniques. Biochemical assays of various tissues from transplanted mice detected CAT enzymatic activity specifically in bone tissues (femur and calvaria), but not in soft tissues (heart, lung, spleen, liver, etc.) Immunohistochemical studies using an antibody against CAT further demonstrated that CAT proteins are expressed primarily in osteoblasts and osteocytes located in bone tissues of transplanted mice.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

- 32 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT:

- (A) NAME: UNIVERSITY OF MASSACHUSETTS
- (B) STREET: 55 LAKE AVENUE, NORTH
- (C) CITY: WORCESTER
- 10 (D) STATE: MASSACHUSETTS
- (E) COUNTRY: US
- (F) POSTAL CODE (ZIP): 01655-0243
- (G) TELEPHONE:
- (H) TELEFAX:

15

- (ii) TITLE OF INVENTION: GENE THERAPY USING BONE MARROW TRANSPLANTS
TRANSFECTED WITH THERAPEUTIC GENES UNDER
THE CONTROL OF TISSUE-SPECIFIC PROMOTERS

20

- (iii) NUMBER OF SEQUENCES: 22

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: LAHIVE & COCKFIELD, LLP
- (B) STREET: 28 STATE STREET
- 25 (C) CITY: BOSTON
- (D) STATE: MASSACHUSETTS
- (E) COUNTRY: US
- (F) ZIP: 02109-1875

30

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

35

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: PCT/US98
- (B) FILING DATE: 05 MARCH 1998
- (C) CLASSIFICATION:

40

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US 60/039,839
- (B) FILING DATE: 06 MARCH 1997

45

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: MANDRAGOURAS, AMY E.
- (B) REGISTRATION NUMBER: 36,207
- (C) REFERENCE/DOCKET NUMBER: UMM-007PC

50

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (617) 227-7400
- (B) TELEFAX: (617) 742-4214

(2) INFORMATION FOR SEQ ID NO:1:

55

- 33 -

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1675 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 496..559

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 816..848

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1025..1096

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1298..1427

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACGGGGCTGA	CAGTAGAAAT	CACAGGCTGT	GAGACAGCTG	GAGCCCAGCT	CTGCTTGAAC	60
CTATTTTAGG	TCTCTGATCC	CCGCTTCCTC	TTTAGACTCC	CCTAGAGCTC	AGCCAGTGCT	120
CAACCTGAGG	CTGGGGGTCT	CTGAGGAAGA	GTGAGTTGGA	GCTGAGGGGT	CTGGGGCTGT	180
CCCCTGAGAG	AGGGGCCAGA	GGCAGTGTCA	AGAGCCGGGC	AGTCTGATTG	TGGCTCACCC	240
TCCATCACTC	CCAGGGGCCC	CTGGCCCAGC	AGCCGCAGCT	CCCAACCACA	ATATCCTCTG	300
GGGTTTGGCC	TACGGAGCTG	GGGCGGATGA	CCCCCAAATA	GCCCTGGCAG	ATTCCCCCTA	360
GACCCGCCCC	CACCATGGTC	AGGCATGCCC	CTCCTCATCG	CTGGGCACAG	CCCAGAGGGT	420
ATAAACAGTG	CTGGAGGCTG	GCGGGGCAGG	CCAGCTGAGT	CCTGAGCAGC	AGCCCAGCGC	480
AGCCACCGAG	ACACC	ATG	AGA	GCC	CTC	ACA
	Met	Arg	Ala	Leu	Thr	Leu
	1		5		10	
GCC	GCA	CTT	TGC	ATC	GCT	GGC
Ala	Ala	Leu	Cys	Ile	Ala	Gln
	15		20			
TCAGGCCGCA	TTGCAGTGGG	GGCTGAGAGG	AGGAAGCACC	ATGGCCCACC	TCTTCTCACC	639
CCTTTGGCTG	GCAGTCCCTT	TGCAGTCTAA	CCACCTTGTT	GCAGGCTCAA	TCCATTGCCC	699
CCAGCTCTGC	CCTTGACAGAG	GGAGAGGAGG	GAAGAGCAAG	CTGCCCGAGA	CGCAGGGGAA	759

- 34 -

	GGAGGATGAG GGCCCTGGGG ATGAGCTGGG GTGAACCAGG CTCCCTTTCC TTTGCA	815
5	GGT GCG AAG CCC AGC GGT GCA GAG TCC AGC AAA GGTGCAGGTA TGAGGATGGA Gly Ala Lys Pro Ser Gly Ala Glu Ser Ser Lys 1 5 10	868
	CCTGATGGGT TCCTGGACCC TCCCCTCTCA CCCTGGTCCC TCAGTCTCAT TCCCCACTC	928
10	CTGCCACCTC CTGTCTGGCC ATCAGGAAGG CCAGCCTGCT CCCCACCTGA TCCTCCCAA	988
	CCCAGAGCCA CCTGATGCCT GCCCCTCTGC TCCACA GCC TTT GTG TCC AAG CAG Ala Phe Val Ser Lys Gln 1 5	1042
15	GAG GGC AGC GAG GTA GTG AAG AGA CCC AGG CGC TAC CTG TAT CAA TGG Glu Gly Ser Glu Val Val Lys Arg Pro Arg Arg Tyr Leu Tyr Gln Trp 10 15 20	1090
20	CTG GGG TGAGAGAAAA GGCAGAGCTG GGCCAAGGCC CTGCCTCTCC GGGATGGTCT Leu Gly	1146
	GTGGGGGAGC TGCAGCAGGG AGTGGCCTCT CTGGGTTGTG GTGGGGGTAC AGGCAGCCTG	1206
25	CCCTGGTGGG CACCCTGGAG CCCCATGTGT AGGGAGAGGA GGGATGGGCA TTTTGCACGG	1266
	GGGCTGATGC CACCACGTCG GGTGTCTCAG A GCC CCA GTC CCC TAC CCG GAT Ala Pro Val Pro Tyr Pro Asp 1 5	1318
30	CCC CTG GAG CCC AGG AGG GAG GTG TGT GAG CTC AAT CCG GAC TGT GAC Pro Leu Glu Pro Arg Arg Glu Val Cys Glu Leu Asn Pro Asp Cys Asp 10 15 20	1366
35	GAG TTG GCT GAC CAC ATC GGC TTT CAG GAG GCC TAT CGG CGC TTC TAC Glu Leu Ala Asp His Ile Gly Phe Gln Glu Ala Tyr Arg Arg Phe Tyr 25 30 35	1414
40	GGC CCG GTC TAG G GTGTCGCTCT GCTGGCCTGG CCGGCAACCC CAGTTCTGCT Gly Pro Val 40	1467
	CCTCTCCAGG CACCCTTCTT TCCTCTTCCC CTTGCCCTTG CCCTGACCTC CCAGCCCTAT	1527
45	GGATGTGGGG TCCCCATCAT CCCAGCTGCT CCCAAATAAA CTCCAGAAGA GGAATCTGTG	1587
	GGCCTGTGAG TCTGTCCAGT TTATGGAGTG TGGGAGGGAG GTGTCAGGAG GATGGGGGTG	1647
50	AGGAGGTTTT ACCTTCTTCA GTTCTAGA	1675

(2) INFORMATION FOR SEQ ID NO:2:

- 35 -

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

10 Met Arg Ala Leu Thr Leu Leu Ala Leu Leu Ala Leu Ala Ala Leu Cys
 1 5 10 15

Ile Ala Gly Gln Ala
 20

15

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 11 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: protein

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Gly Ala Lys Pro Ser Gly Ala Glu Ser Ser Lys
 1 5 10

30

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

40

Ala Phe Val Ser Lys Gln Glu Gly Ser Glu Val Val Lys Arg Pro Arg
 1 5 10 15

45

Arg Tyr Leu Tyr Gln Trp Leu Gly
 20

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: protein

55

- 36 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ala Pro Val Pro Tyr Pro Asp Pro Leu Glu Pro Arg Arg Glu Val Cys
 1 5 10 15
 5 Glu Leu Asn Pro Asp Cys Asp Glu Leu Ala Asp His Ile Gly Phe Gln
 20 25 30
 10 Glu Ala Tyr Arg Arg Phe Tyr Gly Pro Val *
 35 40

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 1557 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 662..725
 25

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 867..899

30 (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1039..1098

(ix) FEATURE:
 35 (A) NAME/KEY: CDS
 (B) LOCATION: 1303..1362

(ix) FEATURE:
 40 (A) NAME/KEY: CDS
 (B) LOCATION: 1364..1436

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCAACCACAA GAAATGCCCT ACAACCGGAT CTTATGGAGG CATTTTCTCA ATTGAGGTTT 60
 45 TCTCCTTCCA AGAAGACATA AAATAACCA GGACATCACC CCCACACACA CACACACACA 120
 CCCACTGGAT GAGCAGAGCT CCCTGAACTG GGCAAATGAG GACATTACTG AACACTCCCT 180
 50 CCCTGGGGTT TGCCTCCCGC TCTCAGGGGC AGACACTGAA AATCACAGGC TATGAGAGTT 240
 GGAGCCCAGG TTATCCCAA CCGATTTTAG ACTTCTGTAC CATGTCTAGG CATGCATAGG 300
 GTTCTTGTCT CTAGGGCGAC CCAGTGCTCC AGCTGAGGCT GAGAGAGAGA GAGCACACAG 360
 55

- 37 -

	TAGGAGTGGT GGAGCAGCCC CTCAGGGAAG AGGTCTGGGG CCATGTCAGA GCCTGGCAGT	420
	CTCCGATTGT GGCTCTCGT CCACTCCCAG AGCCTTGCCC AGGCAGCTGC AATCACCAAC	480
5	CACAGCATCC TTTGGGTTTG ACCCACTGAG CACATGACCC CCAATTAGTC CTGGCAGCAT	540
	CCCCTGCTCC TCCTGCTTAC ATCAGAGAGC ACAGAGTAGC CGATATAAAT GCTACTGGAT	600
10	GCTGGAGGGT GCAGAACAGA CAAGTCCCAC ACAGCAGCTT GGTGCACACC TAGCAGACAC	660
	C ATG AGG ACC CTC TCT CTG CTC ACT CTG CTG GCC CTG GCT GCG CTC	706
	Met Arg Thr Leu Ser Leu Leu Thr Leu Leu Ala Leu Ala Ala Leu	
	1 5 10 15	
15	TGT CTC TCT GAC CTC ACA G GTATGTGTCC TCCTGGTTCA TTTCTTTGGG	755
	Cys Leu Ser Asp Leu Thr	
	20	
20	TAACTACCTC CTGAAGGTCT CACAATCTGC TTTGGGATGG CAGAGGGGAA GGGACAACAC	815
	ATGAGGGAGA CAGCAGGGAG GAAACAGAAC TAACTACCTG TTTGCTTTAC A GAT GCC	872
	Asp Ala	
	1	
25	AAG CCC AGC GGC CCT GAG TCT GAC AAA GGTACTAGCA GGAAGCCTGG	919
	Lys Pro Ser Gly Pro Glu Ser Asp Lys	
	5 10	
30	CAGGGCCTCG GCTTGGCCTC ACCCTGTCCC CTAAGCCCCC AAATCCCCTT GCCTTCTGCC	979
	TGTGTCCCAC TTTTCCTCAC TGAAGTCTGAG ATTACCTGAC CTTGTGTGTC TTCTCCACA	1038
	GCC TTC ATG TCC AAG CAG GAG GGC AAT AAG GTA GTG AAC AGA CTC CGG	1086
	Ala Phe Met Ser Lys Gln Glu Gly Asn Lys Val Val Asn Arg Leu Arg	
35	1 5 10 15	
	CGC TAC CTT GGG TAAGTGCCAG AGCCCTTAGC CTTCCATATT GGTAGGGAGG	1138
	Arg Tyr Leu Gly	
	20	
40	AGTTGTGCTG GGGTGGTTTC TGTGACCCGC AGAGGCTACA CGTGCAGGTC AATCCCCATG	1198
	TCCAGGACCC TGGAGCCTCT TGTACAGTGT GGGAAGAGGG TGTGTGTACC CCGTGTATAT	1258
45	TAATGCCACT GTGTGTTGGT TGATGTTACT TTATACTTCT CAGA GCC TCA GTC CCC	1314
	Ala Ser Val Pro	
	1	
50	AGC CCA GAT CCC CTG GAG CCC ACC CGG GAG CAG TGT GAG CTT AAC CCC	1362
	Ser Pro Asp Pro Leu Glu Pro Thr Arg Glu Gln Cys Glu Leu Asn Pro	
	5 10 15 20	
	T GCT TGT GAC GAG CTA TCA GAC CAG TAT GGC TTG AAG ACC GCC TAC	1408
	Ala Cys Asp Glu Leu Ser Asp Gln Tyr Gly Leu Lys Thr Ala Tyr	
55	1 5 10 15	

- 38 -

AAA CGC ATC TAC GGT ATC ACT ATT TAG G ACCTGTGCTG CCCTAAAGCC 1456
Lys Arg Ile Tyr Gly Ile Thr Ile
20

5

AAACTCTGGC AGCTCGGCTT TGGCTGCTCT CCGGGACTTG ATCCTCCCTG TCCTCTCTCT 1516
CTGCCCTGCA AGTATGGATG TCACAGCAGC TCCAAAATAA A 1557

10 (2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

15

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

20

Met Arg Thr Leu Ser Leu Leu Thr Leu Leu Ala Leu Ala Ala Leu Cys
1 5 10 15

25

Leu Ser Asp Leu Thr
20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

30

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Asp Ala Lys Pro Ser Gly Pro Glu Ser Asp Lys
1 5 10

40

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

45

(A) LENGTH: 20 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ala Phe Met Ser Lys Gln Glu Gly Asn Lys Val Val Asn Arg Leu Arg
1 5 10 15

- 39 -

Arg Tyr Leu Gly
20

(2) INFORMATION FOR SEQ ID NO:10:

5

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

15 Ala Ser Val Pro Ser Pro Asp Pro Leu Glu Pro Thr Arg Glu Gln Cys
1 5 10 15

Glu Leu Asn Pro
20

20

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

30

Ala Cys Asp Glu Leu Ser Asp Gln Tyr Gly Leu Lys Thr Ala Tyr Lys
1 5 10 15

35 Arg Ile Tyr Gly Ile Thr Ile *
20

(2) INFORMATION FOR SEQ ID NO:12:

40

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1848 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

50

- (A) NAME/KEY: CDS
- (B) LOCATION: 942..1005

(ix) FEATURE:

55

- (A) NAME/KEY: CDS
- (B) LOCATION: 1147..1179

- 40 -

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1322..1381

5 (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1588..1647

10 (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1649..1721

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

15	TAAGAGGGAT TTTTGGCGTT GCTGCAGCCA TTGTCACTTC CTTAGCAATC TCTGCCACCG	60
	CAGCCACGGC TGCATATCCC TTACAACTG CTGCGGCTGT GAATAATTTA TCCGCAGTGT	120
	TGCAGAGGTT TTGGATTGCT GACAGCTGTC CCCCTCTTGA AAGAGTGGGC TGGAATTGGT	180
20	GGCATTGCCC TGTGTCCGAT ATTAACAGCA TTCTGGCCTT CTGGTGCACC TGCCGCATGC	240
	AGAGTCGTCA GCGCAGAGCC CAAGCCTTGA TGATACAGGC TTTTGCGGGT GTGAAAACAG	300
25	GACAATCTCC TCAACTGTGG CTTGGCATGC TAGAGAAGTA GTCAATGACG GGTAAGACTC	360
	CCTGGGCGTG TCACCAACCT AAGACAGGGA TCAAACCAAT GTTGTTTGTC TCCCGAGGAC	420
	AGGTAAGGGG CATTGCTGCA GGGGGCAACC TAAGACAGGC ATTCTCTCTG CCAATAAGTA	480
30	AAGACAGGGA GATGTGAGGA GCGGGTGTGG CAGCAGTCCC AAGATGGCGC CCGGGACTGC	540
	AGCCAAGTCT TATGACTTTC ACCTGACTTC CTCATACACC TGAAAATAAG CCACGACAAT	600
35	TGTGAGACTG CGCAGGTGCA ACATGATGCA AGATCAGACC ATATGACAAG TGATGATTCT	660
	GGCCAATGGA CTGCTGTTCC GTGGACAGGG CAGATGGGGC ATGGTTCAGA GGTATATATA	720
	GGGATTGCCA TTGGGGGAGG AGATACAGAC AGAGAGATTC CTGTGGAGAG ACTTATTCTT	780
40	GGAGAGAGAC TCCTGCATGC ATGTTGAAAG GTTCCTGAAT AACTGCTTT GAGAAGAACA	840
	TAGTGTCGTC GTTTCCTTCT GCTGGTCAGA ATCAGAGGCA ACAGGAGGGT GCAGAACAGA	900
45	CAAGTCCAC ACAGCAGCTT GGTGCACACC TAGCAGACAC C ATG AGG ACC CTC	953
	Met Arg Thr Leu	
	1	
	TCT CTG CTC ACT CTG CTG GCC CTG GCT GCG CTC TGT CTC TCT GAC CTC	1001
50	Ser Leu Leu Thr Leu Leu Ala Leu Ala Ala Leu Cys Leu Ser Asp Leu	
	5 10 15 20	
	ACA G GTATGTGTCC TCCTGGTTCA TTTCTTTGGG TGAAGGTCT	1055
55	Thr	

- 41 -

CACAATCTGC TTTGGGATGG CAGAGGGGAA GGGACAACAC ATGAGGGAGA CAGCAGGGAG 1115
 5 GAAACAGAAC TAACTACCTG TTTGCTTTAC A GAT GCC ACG CCC ACC GGC CCT 1167
 Asp Ala Thr Pro Thr Gly Pro
 1 5
 GAG TCT GAC AAA GGTACTGGCA GAAAGCCTGG CAGGGCCTCG GCTTGGCCAC 1219
 10 Glu Ser Asp Lys
 10
 ACCCTGTCCC CTAAGCCCCC AAATCCCCTT GCCTTCTGCC TGGGTGTCCC CACTTTTCCT 1279
 15 CACTGAACTC AGAATTACCT GACCTTGTGT GTCTTCTCCA CA GCC TTC ATG TCC 1333
 Ala Phe Met Ser
 1
 AAG CAG GAG GGC AAT AAG GTA GTG AAC AGA CTC CGG CGC TAC CTT GGG 1381
 20 Lys Gln Glu Gly Asn Lys Val Val Asn Arg Leu Arg Arg Tyr Leu Gly
 5 10 15 20
 TAAGTGCCAG AGCCCTTAGC CTTCCATATT GGTAGGGAGG AGTTGTTCTG GGGTAGTCTC 1441
 25 TATGACCCGC AGAGGGCTAC ACGTGCAGGT CAATCCCCAG GTCCAGGACC CTGGAGCCTC 1501
 TTGTACAGTG TGGGAAGAGG ATGTGTGTAC CCCTGTGTAT ATTAATGCCA CTGTGTGTTG 1561
 GTTGATGTTA CTTTATGCTT CTCAGA GCT TCA GTC CCC AGC CCA GAT CCC CTG 1614
 30 Ala Ser Val Pro Ser Pro Asp Pro Leu
 1 5
 GAG CCC ACC CGG GAG CTA TGT GAG CTT GAC CCC T GCT TGT GAC GAG 1660
 Glu Pro Thr Arg Glu Leu Cys Glu Leu Asp Pro Ala Cys Asp Glu
 10 15 20 1
 35 CTA TCA AAC CAG TAT GGC TTA AAG ACC GCC TAC AGA CGC ATC TAC GGT 1708
 Leu Ser Asn Gln Tyr Gly Leu Lys Thr Ala Tyr Arg Arg Ile Tyr Gly
 5 10 15 20
 40 ATC ACT ATT TAG G ACCTGTAGCT GCCCTAAAGC CAAACTCTGG CAGCTCGGCT 1761
 Ile Thr Ile *
 TTGGCTGCTC TCCCAGAACT TGACCCTCCC CTGTCCCTCT CTCTCTGCCC CTGCAAGTAT 1821
 45 GGATGTCACA GCACGCTCCA AAATAAA 1848

(2) INFORMATION FOR SEQ ID NO:13:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 55 (ii) MOLECULE TYPE: protein

- 42 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

5 Met Arg Thr Leu Ser Leu Leu Thr Leu Leu Ala Leu Ala Ala Leu Cys
 1 5 10 15
 Leu Ser Asp Leu Thr
 20

10 (2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 11 amino acids
 (B) TYPE: amino acid
 15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

20 Asp Ala Thr Pro Thr Gly Pro Glu Ser Asp Lys
 1 5 10

25 (2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 amino acids
 (B) TYPE: amino acid
 30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

35 Ala Phe Met Ser Lys Gln Glu Gly Asn Lys Val Val Asn Arg Leu Arg
 1 5 10 15

Arg Tyr Leu Gly
 20

40 (2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 amino acids
 45 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Ala Ser Val Pro Ser Pro Asp Pro Leu Glu Pro Thr Arg Glu Leu Cys
 1 5 10 15

- 43 -

Glu Leu Asp Pro
20

(2) INFORMATION FOR SEQ ID NO:17:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

15 Ala Cys Asp Glu Leu Ser Asn Gln Tyr Gly Leu Lys Thr Ala Tyr Arg
1 5 10 15

Arg Ile Tyr Gly Ile Thr Ile *
20

20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2218 base pairs

25

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

30

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1131..1194

35

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1342..1374

(ix) FEATURE:

40

(A) NAME/KEY: CDS

(B) LOCATION: 1511..1582

(ix) FEATURE:

45

(A) NAME/KEY: CDS

(B) LOCATION: 1782..1914

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

50 AATTCAAGAA CAACCTTCAC TTTAATATTA TTGATAACAT TTAGTTTCTG GATATCAAGT 60

GGGCTCCTAC TAGATCATGC CAGGTCACCA AATACCAGTT TAATCAAGGA AGGAAGAAAG 120

AAAAGGAACA GTAAATAAAA AGAGTGACAA GTACTAAGAA TTGCAAAAAG GACATCCAAG 180

55 GTGTTTGGAC CAAAGAAGCG GGCCTTGGGC TCTCAGTACT CATACTGGGC CCAAGAGAC 240

- 44 -

	CATGGCCATT GCTCTGAAAT ACGTAACCCG CAGTTTTTCT TCCTTGTCCA TCAGGGGTCC	300
5	CAGGCATCTT GAGCTTATGT GGGGTGTCTC TGACACAAGC AGGGCTAGAA CCTTGAGGCA	360
	GAAGCTCGAA GCCATGGAGA AGTGCAGCTT ACTGGCTTGT TCATCATGGC TTGCTCAGCC	420
	TGCTGTCTTG TAGAACTCAA GACCAAGGCC CAGGGGTGGC ACCACTCACA ATGGGTGGG	480
10	CTCTCCCAT CAAAACCACT AAATAAGAAA TGCCCTACAA CCGGATCTTA TGGAGGCATT	540
	TTCTCAATTG AGGCTTTCTT CTTTCAAGTT GACAGAAAAC TAAGCAGGAC ACTTCCCCAC	600
15	CCTTCTCCCC ACTGGATGAG CGGAGCTGCC CTGCACTGGG TGAATGAGGA CATTACTGAC	660
	CGCTCCTTCC TGGGGTTTGG CTCCTGCTCT CACGGCAGAC ACTGAAAATC ACAGGCTATG	720
	CGAGTTGGAG CCAGTTTGTC CCAAACCGAT TTTAGATCTC TGTACCCTCT CTAGGCTATG	780
20	CTCTAAGGAG CTCGTCTCTA GGGCCAGCCA GTGCTCCAGC TGAGGCTGAG AGAGATGGCA	840
	CACAGTAGGG GTGCTGGAGC AGCCCCTCCG GGAAGAGGTC TGGGGCCATG TCAGGACCCG	900
25	GCAGCCTCTG ATTGTGTCCT ACCCTCCCCT TCCAGGCCTT CGCCCCGGCA GCTGCAGTCA	960
	CCAACCACAG CATCCTTTGG GTTTGACCTA TTGCGCACAT GACCCCCAAT TAGTCCTGGC	1020
	AGCATCTCCT GCCCCTCCTG CTTACATTAG GGGCTCAGGC AGTGGATATA AAAGGTATTG	1080
30	CAGAACAGAC AAGTCCCACA CAGCAACTCG GTGCAGACCT AGCAGACACC ATG AGG	1136
		Met Arg
		1
35	ACC CTC TCT CTG CTC ACT CTG CTG GCC CTG ACT GCA TTC TGC CTC TCT	1184
	Thr Leu Ser Leu Leu Thr Leu Leu Ala Leu Thr Ala Phe Cys Leu Ser	
	5 10 15	
40	GAC CTG GCA G GTATGTGCCC TCCTGGTTCA TTTCTTTGGA TAGCCACCCT	1234
	Asp Leu Ala	
	20	
	CCTGAAGGTC TCGGGATCTG CTTTGGGATG GCAGAAGGGA AGGGACAAAC TACCAGAGAG	1294
45	ACAGACAGCA GGGAGGGAAC AGAACTAACT GCACTGTTTT CTTTACA GGT GCA AAG	1350
		Gly Ala Lys
		1
50	CCC AGC GAC TCT GAG TCT GAC AAA GGTACTGACA GGGAGCCTGG CAGAACCTCA	1404
	Pro Ser Asp Ser Glu Ser Asp Lys	
	5 10	
	GCCTTGCCCC TCACCCTGTC CCTAAGCCCC CAGACCCCTC TGCCTTCTGT CTGGGTGTCC	1464

- 45 -

CACTTTCCTC CAAACCAGAA TACCTGATCC TGTGCCTCTG TTCACA GCC TTC ATG 1519
Ala Phe Met
1

5 TCC AAG CAG GAG GGC AGT AAG GTG GTG AAT AGA CTC CGG CGC TAC CTC 1567
Ser Lys Gln Glu Gly Ser Lys Val Val Asn Arg Leu Arg Arg Tyr Leu
5 10 15

10 AAC AAT GGA CTT GGG TGAGTGGGGA GGAAGGCGCT GCCAGGGCCC TTTGCTTTCC 1622
Asn Asn Gly Leu Gly
20

ATATTGGTAG GGAGGAGTTG TGCTGGGTGG TCTCTATGAA CTGCCGAGGG TCCCATGCCT 1682

15 AGGACCCTGG AGCCTCTTGT GCAGTGGGAA GAGTGTGTGT ATACCCCGTG TTGGTTAATG 1742

CCACTGCGTA TTGGTTGACG CTACTATATG CTTCTCAGA GCC CCA GCC CCC TAC 1796
Ala Pro Ala Pro Tyr
1 5

20 CCA GAT CCC CTG GAG CCT CAC AGG GAG GTG TGT GAG CTC AAC CCC AAT 1844
Pro Asp Pro Leu Glu Pro His Arg Glu Val Cys Glu Leu Asn Pro Asn
10 15 20

25 TGT GAC GAG CTA GCG GAC CAC ATT GGC TTC CAG GAC GCC TAC AAG CGC 1892
Cys Asp Glu Leu Ala Asp His Ile Gly Phe Gln Asp Ala Tyr Lys Arg
25 30 35

30 ATC TAT GGC ACC ACC GTT TAG G GCATGTGTTG CCCTGGAGCC CAACGCAGCT 1944
Ile Tyr Gly Thr Thr Val
40

TCAGCTTTTG GCTACTCTCC AGGACTCGAC CCTCCCTGTT CCCTCTCTCT GCCTCGAAAG 2004

35 TATGGACGGC ACAGCTGCTC CAAAATAAAG TCCAGATGAG GAACGGTTGG GCTCGAGTCT 2064

GTCCATTGTG GTGGTGGTAG GGGTTGGAGA AGGCTTCATG TTCATCCCAG ACAGCCCCAG 2124

40 TCCGGTTACT TTGCGTCTTT ATTATACTGC TCTGTGATGG ATCGCACCAG AGGTCCCACA 2184

GCCCGCCTCC CACAGTGATT CCACCACATC ATAT 2218

(2) INFORMATION FOR SEQ ID NO:19:

45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Arg Thr Leu Ser Leu Leu Thr Leu Leu Ala Leu Thr Ala Phe Cys
55 1 5 10 15

- 46 -

Leu Ser Asp Leu Ala
20

5 (2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

15

Gly Ala Lys Pro Ser Asp Ser Glu Ser Asp Lys
1 5 10

20 (2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

30 Ala Phe Met Ser Lys Gln Glu Gly Ser Lys Val Val Asn Arg Leu Arg
1 5 10 15

Arg Tyr Leu Asn Asn Gly Leu Gly
20

35

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
40 (B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Ala Pro Ala Pro Tyr Pro Asp Pro Leu Glu Pro His Arg Glu Val Cys
1 5 10 15

50 Glu Leu Asn Pro Asn Cys Asp Glu Leu Ala Asp His Ile Gly Phe Gln
20 25 30

Asp Ala Tyr Lys Arg Ile Tyr Gly Thr Thr Val *
35 40

55

CLAIMS

1. A method for expressing an exogenous gene in a differentiated cell,
comprising contacting a pluripotent stem cell capable of maturing into differentiated cell
5 with
a nucleic acid comprising an exogenous gene operably linked to a
regulatory element capable of controlling expression of the exogenous gene in a
differentiated cell, to thereby produce a transduced pluripotent stem cell capable of
maturing into a differentiated cell, which differentiated cell expresses the exogenous
10 gene.
2. The method of claim 1, wherein the differentiated cell is in a tissue
3. The method of claim 2, wherein the tissue is skeletal tissue.
15
4. The method of claim 3, wherein the skeletal tissue is bone.
5. The method of claim 3, wherein the skeletal tissue is cartilage.
- 20 6. The method of claim 4, wherein the differentiated cell is a bone cell.
7. The method of claim 6, wherein the bone cell is an osteoblast.
8. The method of claim 6, wherein the bone cell is an osteocyte.
25
9. The method of claim 3, wherein the regulatory element comprises at least
one promoter.
10. The method of claim 9, wherein the regulatory element comprises an
30 osteocalcin proximal promoter.
11. The method of claim 10, wherein the the osteocalcin proximal promoter
comprises a nucleotide sequence from about -40 to about -200 upstream from the
transcription initiation site of SEQ ID NO:1.
35
12. The method of claim 9, wherein the regulatory element comprises an
osteocalcin distal promoter.

13. The method of claim 12, wherein the osteocalcin distal promoter comprises from about -400 to about -600 upstream from the transcription initiation site of SEQ ID NO:1.

5

14. The method of claim 9, wherein the regulatory element comprises an osteocalcin far distal promoter.

15. The method of claim 14, wherein the osteocalcin far distal promoter comprises from about -600 to about -700 upstream from the transcription initiation site of SEQ ID NO:1.

10

16. The method of claim 9, wherein the regulatory element binds homeodomain containing transcription factors and osteocalcin binding proteins.

15

17. The method of claim 16, wherein the regulatory element comprises from about nucleotides -80 to about -100 upstream from the transcription initiation site of SEQ ID NO:1.

20

18. The method of claim 9, wherein the regulatory element binds the AML family of transcription factors.

19. The method of claim 18, wherein the regulatory element comprises from about nucleotides -70 to about -170 upstream from the transcription initiation site of SEQ ID NO:1.

25

20. The method of claim 3, wherein the regulatory element comprises at least two osteocalcin transcriptional control elements.

30

21. The method of claim 20, wherein the osteocalcin transcriptional control elements are an osteocalcin box I and an osteocalcin box II.

22. The method of claim 21, wherein the osteocalcin box I and the osteocalcin box II are a portion of a nucleic acid molecule having a nucleotide sequence shown in SEQ ID NO: 1.

35

- 49 -

23. The method of claim 20, wherein the regulatory elements comprise the osteocalcin distal promoter and the osteocalcin proximal promoter.

24. The method of claim 10, wherein the regulatory element further
5 comprises a distal promoter of an osteocalcin gene.

25. The method of claim 24, wherein the regulatory element further comprises a far distal promoter of an osteocalcin gene.

10 26. The method of claim 3, wherein the regulatory element comprises at least about 1.7 kb of the sequence shown in SEQ ID NO: 1.

27. The method of claim 3, wherein the regulatory element comprises at least about 0.35 kb of the sequence shown in SEQ ID NO: 1.
15

28. The method of claim 9, wherein the regulatory element is inducible.

29. The method of claim 28, wherein the regulatory element is a promoter responsive to vitamin D.
20

30. The method of claim 28, wherein the regulatory element is a promoter responsive to a factor selected from the group consisting of a steroid hormone, a growth factor, a cytokine, a morphogenic protein, a polypeptide hormone, signal transduction factors, and intermediates.
25

31. The method of claim 3, wherein the exogenous gene encodes a therapeutic protein useful for treating a disease.

32. The method of claim 31, wherein the therapeutic protein is selected from
30 the group consisting of steroid hormones, steroid hormone receptors, growth factors, cytokines, morphogenic proteins, polypeptide hormones, polypeptide chemotherapeutic agents, signal transduction factors and intermediates.

33. The method of claim 32, wherein the disease is a skeletal disease selected
35 from the group consisting of: osteoporosis, osteopenia, and osteosarcoma.

- 50 -

34. The method of claim 32, wherein the disease is a primary malignancy or metastasis.

5 35. The method of claim 3, wherein the exogenous gene encodes a growth hormone.

10 36. The method of claim 3, further comprising contacting the bone marrow with a second nucleic acid comprising a second exogenous gene operably linked to a second promoter capable of controlling expression of the exogenous gene in a differentiated cell.

37. The method of claim 1, further comprising administering the transduced population of pluripotent stem cells to a subject.

15 38. The method of claim 37, wherein the differentiation occurs *in vivo*.

39. The method of claim 1, wherein the stem cell is a bone marrow cell obtained from a subject.

20 40. The method of claim 39, wherein the transduced population of stem cells is administered to the subject from whom the bone marrow was obtained.

1/11

1 acggggctga cagtagaaat cacaggctgt gagacagctg gagcccagct ctgcttgaac
 61 ctattttagg tctctgatcc ccgcttcctc tttagactcc cctagagctc agccagtgt
 121 caacctgagg ctgggggtct ctgaggaaga gtgagttgga gctgaggggt ctggggctgt
 181 cccctgagag agggggccaga ggcagtgtca agagccgggc agtctgattg tggctcacc
 241 tccatcactc ccaggggccc ctggcccagc agccgcagct cccaaccaca atatcctctg
 301 gggtttgcc tacggagctg gggcggatga ccccaaata gccctggcag attcccccta
 361 gacccgccc caccatggtc aggcagccc ctctcatcg ctgggcacag ccagaggggt
 421 ataaacagt ctggaggctg gcggggcagg ccagctgagt cctgagcagc agcccagcgc
 481 agccaccgag acaccatgag agccctcaca ctctcgccc tattggccct ggccgcactt
 541 tgcctgctg gccaggcagg tgagtgcgcc cactccccct caggccgcat tgcagtggg
 601 gctgagagga ggaagcacca tggcccacct ctctcacc ctttgctgg cagtccctt
 661 gcagttaac cacttgttg caggctcaat ccatgtccc cagctctgcc ctgacagagg
 721 gagaggagg aagagcaagc tgcccagac gcagggaag gaggatgagg gccctgggga
 781 tgagctggg tgaaccaggc tcccttctc ttgcagggtc gaagcccagc ggtgcagagt
 841 ccagcaaagg tgcaggtatg aggatggacc tgatgggttc ctggaccctc cctctcacc
 901 ctggccctc agtctattc cccactcct gccacctct gtctggccat caggaaaggcc
 961 agcctgctc ccacctgatc ctccaaacc cagagccacc tgatgcctgc ccctctgct
 1021 cacagcctt gtgtccaagc aggagggcag cgaggtagt aagagacca ggcgtacct
 1081 gtatcaatg ctgggtgag agaaaaggca gagctgggccc aaggccctgc ctctccggga
 1141 tggctgttg gggagctgca gcaggagtg gcctctctg gttgtgttg gggtaggc
 1201 agcctgcct ggtgggcacc ctggagccc atgttaggg agaggaggga tggcatitt
 1261 gcacgggggc tgatgccacc acgtcgggtg tctcagagcc ccagtcctt acccgatcc
 1321 cctggagccc aggaggagg tgtgtgagt caatccggac tgtacgagt tggctgacca
 1381 catcggttt caggaggcct atcggcgtt ctacggccc gtctaggggt tgcctctgt
 1441 ggctggccg gcaacccag ttctgtcct ctccaggcac cttcttcc tcttccctt
 1501 gccctgccc tgacctcca gccctatgga tgtgggtcc ccatcatccc agctgtccc
 1561 aaataaact cagaagagga atctgtggc ctgtgagtct gtccagttta tggagtgtg
 1621 gagggaggtg tcaggaggat gggggtagg aggtttacc ttctcagtt ctaga

FIGURE 1

2/11

1 ccaaccacaa gaaatgcctt acaaccggat cttatggagg cattttctca
 attgaggttt
 61 tctccttcca agaagacata aaactaacca ggacatcacc cccacacaca
 cacacacaca
 121 ccactggat gagcagagct ccctgaactg ggcaaatgag gacattactg
aacactccct
 181 ccctggggtt tgcgtcccg ctcaggggc agacactgaa aatcacaggc
 tatgagagtt
 241 ggagcccagg ttatcccaaa ccgatttttag acttctgtac catgtctagg
 catgcatagg
 301 gttcttgtct ctagggcgac ccagtgtcc agctgaggt gagagagaga
 gagcacacag
 361 taggagtggg ggagcagccc ctcaggggag aggtctgggg ccattgtcaga
 gcctggcagt
 421 ctccgattgt ggccctctgt ccactcccag agccttgccc aggcagctgc
 aatcaccaac
 481 cacagcatcc tttgggtttg acccactgag cacatgaccc ccaattagtc
ctggcagcat
 541 cccctgtccc tcctgcttac atcagagagc acagagtagc cgatataaat
 gctactggat
 601 gctggagggg gcagaacaga caagtccac acagcagctt ggtgcacacc
 tagcagacac
 661 catgaggacc ctctctctgc tctactctgt ggccctgggt gcgtctctgc
 tctctgacct
 721 cacaggtatg tgctctctgt gttcatttct ttgggtaact acctcctgaa
 ggtctcacia
 781 tctgctttgg gatggcagag ggggaaggag aacacatgag ggagacagca
 gggaggaaac
 841 agaactaact acctgtttgc ttacagatg ccaagcccag cggccctgag
 tctgacaaag
 901 gtactagcag gaagcctggc agggcctcgg cttggcctca ccctgtcccc
 taagccccc
 961 aatccccttg ccttctgcct gtgtccact tttctcact gaactcagaa
 ttacctgacc
 1021 ttgtgtgtct tctccacagc cttcatgtcc aagcaggagg gcaataaggt
 agtgaacaga
 1081 ctccggcgct acctgggta agtgccagag cccttagcct tccatattgg
 tagggaggag
 1141 ttgtgctggg gtggtttctg tgaccgcag aggtacacg tgcaggtcaa
 tccccatgtc
 1201 caggaccctg gagcctcttg tacagtgtgg gaagaggggt tgtgtacccc
 gtgtatatta
 1261 atgccactgt gtgttggttg atgttacttt atacttctca gagcctcagt
 cccagccca
 1321 gatcccctgg agcccacccg ggagcagtg gagcttaacc cctgcttggtg
 acgagctatc
 1381 agaccagtat ggcttgaga ccgcctacaa acgcatctac ggtatcacta
 tttaggacct
 1441 gtgctgccct aaagccaaac tctggcagct cggctttggc tgcctctccgg
 gacttgatcc
 1501 tccctgtcct ctctctctgc cctgcaagta tggatgtcac agcagctcca
 aaataaa

FIGURE 2A

3/11

```

1 taagagggat ttttggcggt gctgcagcca ttgtcacttc cttagcaatc
tctgccaccg
61 cagccacggc tgcatatccc ttacaaactg ctgcggctgt gaataattta
tccgcagtgt
121 tgcagagggt ttggattgct gacagctgtc cccctcttga aagagtgggc
tggaattggt
181 ggcattgccc tgtgtccgat attaacagca ttctggcctt ctggtgcacc
tgccgcacgc
241 agagtcgtca gcgcagagcc caagccttga tgatacaggc ttttgcgggt
gtgaaaacag
301 gacaatctcc tcaactgtgg cttggcatgc tagagaagta gtcaatgacg
ggtaagactc
361 cctgggcggtg tcaccaacct aagacaggga tcaaaccaat gttgtttgtc
tcccaggac
421 aggttaagggg cattgctgca gggggcaacc taagacaggc attctctctg
ccaataagta
481 aagacaggga gatgtgagga gcgggtgtgg cagcagtcctc aagatggcgc
ccgggactgc
541 agccaagtct tatgactttc acctgacttc ctcatacacc tgaaaataag
ccacgacaat
601 tgtgagactg cgcaggtgca acatgatgca agatcagacc atatgacaag
tgatgattct
661 ggccaatgga ctgctgttcc gtggacaggg cagatggggc atggttcaga
ggttatataa
721 gggattgcca ttgggggagg agatacagac agagagattc ctgtggagag
acttattctt
781 ggagagagac tcctgcacgc atgttgaaag gttcctgaat aaactgcttt
gagaagaaca
841 tagtgtcgtc gtttctttct gctggtcaga atcagaggca acaggagggt
gcagaacaga
901 caagtccac acagcagctt ggtgcacacc tagcagacac catgaggacc
ctctctctgc
961 tcactctgct ggccctggct gcgctctgtc tctctgacct cacaggatg
tgtcctcctg
1021 gttcatttct ttgggtgact acctcctgaa ggtctcaca tctgctttgg
gatggcagag
1081 gggaaggggac aacacatgag ggagacagca gggaggaaac agaactaact
acctgtttgc
1141 ttacagatg ccaagccac cgccctgag tctgacaaag gtactggcag
aaagcctggc
1201 agggcctcgg cttggccaca ccctgtcccc taagccccca aatccccctg
ccttctgcct
1261 ggggtgtccc acttttctc actgaactca gaattacctg acctgtgtg
tcttctccac
1321 agccttcacg tccaagcagg agggcaataa ggtagtgaac agactccggc
gctaccttgg
1381 gtaagtgcc aagcccttag ccttccatat tggtagggag gagttgttct
gggtagtct
1441 ctatgaaccg cagagggtta cacgtgcagg tcaatccccca ggtccaggac
cctggagcct
1501 cttgtacagt gtgggaagag gatgtgtgta ccctgtgta tattaatgcc
actgtgtgtt

```

FIGURE 2B

4/11

```
1561 ggttgatggt actttatgct tctcagagct tcagtcccca gccagatcc  
cctggagccc  
1621 acccgggagc tatgtgagct tgaccctgc ttgtgacgag ctatcaaacc  
agtatggctt  
1681 aaagaccgcc tacagacgca tctacggtat cactatttag gacctgtagc  
tgccctaaag  
1741 ccaaactctg gcagctcggc ttggctgct ctcccagaac ttgaccctcc  
cctgtccctc  
1801 tctctctgcc cctgcaagta tggatgtcac agcacgctcc aaaataaa
```

FIGURE 2B, continued

5/11

1 aattcaagaa caaccttcac tttaatatta ttgataacat ttagtttctg
 gatatacaagt
 61 gggctcctac tagatcatgc caggtcacca aataccagtt taatcaagga
 aggaagaaag
 121 aaaaggaaca gtaaataaaa agagtgacaa gtactaagaa ttgcaaaaag
 gacatccaag
 181 gtgttttgac caaagaagcg ggccttgggc tctcagtact catactgggc
 cccaagagac
 241 catggccatt gctctgaaat acgtaaccg cagtttttct tccttgcca
 tcaggggtcc
 301 caggcatctt gagcttatgt ggggtgtctc tgacacaagc agggctagaa
 ccttgaggca
 361 gaagctcgaa gccatggaga agtgcagctt actggcttgt tcatcatggc
 ttgctcagcc
 421 tgctgtcttg tagaactcaa gaccaaggcc caggggtggc accactcaca
 atgggttggg
 481 ctctcccat caaaaccact aaataagaaa tgccctacaa ccgatctta
 tggaggcatt
 541 ttctcaattg aggttttctt ctttcaagtt gacagaaaac taagcaggac
 acttccccac
 601 ccttctcccc actggatgag cggagctgcc ctgcactggg tgaatgagga
 cattactgac
 661 cgctccttcc tggggtttgg ctctgtctc caggcagac actgaaaac
 acaggctatg
 721 cgagttggag ccagtttgtc ccaaaccgat ttagatctc tgtaccctct
 ctaggctatg
 781 ctctaaggag ctctgtctta gggccagcca gtgtccagc tgaggctgag
 agagatggca
 841 cacagtaggg gtgtgggagc agccctccg ggaagaggtc tggggccatg
 tcaggaccg
 901 cgagcctctg attgtgtcct accctccct tccaggcctt cggccggca
 gctgcagtca
 961 ccaaccacag catcctttgg gtttgaccta ttgcgcacat gaccccaat
 tagtcctggc
 1021 agcatctcct gcccctcctg cttacattag gggctcaggc agtggatata
 aaaggtattg
 1081 cagaacagac aagtcacaca cagcaactcg gtgcagacct agcagacacc
 atgaggacc
 1141 tctctctgct cactctgctg gccctgactg cattctgcct ctctgacctg
 gcaggatatg
 1201 gccctcctgg ttcatttctt tggatagcca cctcctgaa ggtctcggga
 tctgctttg
 1261 gatggcagaa gggaaggagc aaactaccag agagacagac agcagggagg
 gaacagaact
 1321 aactgcactg ttttctttac aggtgcaaag ccagcgact ctgagtctga
 caaaggact
 1381 gacagggagc ctggcagaac ctcagccttg cccctcacc tgctccctaag
 ccccgagacc
 1441 -cctctgcctt ctgtctgggt gtcccactt cctccaaacc agaataacctg
 atcctgtgcc

FIGURE 3

6/11

1501 tctgttcaca gccttcatgt ccaagcagga gggcagtaag gtggtgaata
gactccggcg
1561 ctacctcaac aatggacttg ggtgagtggg gaggaaggcg ctgccagggc
cctttgcttt
1621 ccatattggt agggaggagt tgtgctgggt ggtctctatg aactgccgag
ggccccatgc
1681 ctaggaccct ggagcctctt gtgcagtggg aagagtgtgt gtataccccg
tgttggttaa
1741 tgccactgcg tattgggtga cgctactata tgcttctcag agccccagcc
ccctaccag
1801 atcccctgga gcctcacagg gaggtgtgtg agctcaacc caattgtgac
gagctagcgg
1861 accacattgg cttccaggac gcctacaagc gcctctatgg caccaccgtt
tagggcatgt
1921 gttgccctgg agcccaacgc agcttcagct tttggctact ctccaggact
cgaccctccc
1981 tgttccctct ctctgcctcg aaagtatgga cggcacagct gctccaaaat
aaagtccaga
2041 tgaggaacgg ttgggctcga gtctgtccat tgtggtggtg gtaggggttg
gagaaggctt
2101 catgttcac ccagacagcc ccagtccggt tactttgcgt ctttattata
ctgctctgtg
2161 atggatcgca ccagaggtcc cacagccgc ctcccacagt gattccacca
catcatat

FIGURE 3, continued

7/11

Osteocalcin Gene Sequence Comparison

	1				50
mouseCCA	ACCACAAGAA	ATGCCCTACA	ACCGGATCTT
rat	GCTCTCCCCA	TCAAAACCAC	TAAATAAGAA	ATGCCCTACA	ACCGGATCTT
human
	51				100
mouse	ATGGAGGCAT	TTTCTCAATT	GAGGTTTCT	CCTTCCAAGA	AGACATAAAA
rat	ATGGAGGCAT	TTTCTCAATT	GAGGCTTTCT	TCTTTCAAGT	TGACAGAAAA
human
	101				150
mouse	CTAACCAGGA	CATCACCCCC	ACACACACAC	ACACACACCC	ACTGGATGAG
rat	CTAAGCAGGA	CACTTCCC..CAC	CCTTCTCCCC	ACTGGATGAG
human
	151				200
mouse	CAGAGCT.CC	CTGAACTGGG	CAAATGAGGA	CATTACTGAA	CACTCCCTCC
rat	CGGAGCTGCC	CTGCACTGGG	TGAATGAGGA	CATTACTGAC	CGCT.CCTTC
human
	201				250
mouse	CTGGGGTTTG	CGTCCCCTC	TCAGGGCAG	ACACTGAAAA	TCACAGGCTA
rat	CTGGGGTTTG	GCTCCTGCTC	TCA.CGGCAG	ACACTGAAAA	TCACAGGCTA
humanACGGGGCTG	ACAGTAGAAA	TCACAGGCTG
	251				300
mouse	TGAG..AGTT	GGAGCCCAGG	TTATCCCAAA	CCGATTTTAG	ACTTCTGTAC
rat	TGCG..AGTT	GGAG.CCAGT	TTGTCCCAAA	CCGATTTTAG	ATCTCTGTAC
human	TGAGACAGCT	GGAGCCCAGC	TCTGCTTGAA	CCTATTTTAG	GTCTCTGATC
	301				350
mouse	CATGTCTAGG	C..ATGCATA	GGGTTCTTGT	C.TCTAGGGC	GACCCAGTGC
rat	CCTCTCTAGG	CTATGCTCTA	AGGAGCTCGT	C.TCTAGGGC	CAGCCAGTGC
human	CCCGCTTCCT	CTTTAGACTC	CCCTAGAGCT	CAGCCAGTGC	TCAACCTGAG
	351				400
mouse	TCCAGCTGAG	GCTGAGAGAG	AGAGAGCACA	CAGTAGGAGT	GGTGGAGCAG
rat	TCCAGCTGAG	GCT..GAGAG	AGATGGCACA	CAGTAGGGGT	GCTGGAGCAG
human	GCTGGGGGTC	TCTGAGGAAG	AGTGAGTTGG	AGCTGAGGGG	TCTGGGGGCTG
	401				450
mouse	CCCCTCAGGG	AAGAGGTCTG	GGGCCATGTC	AGAGCCTGGC	AGTCTCCGAT
rat	CCCCTCCGGG	AAGAGGTCTG	GGGCCATGTC	AGGACCCGGC	AGCCTCTGAT
human	TCCCCTGAGA	GAGGGGCCAG	AGGCAGTGTC	AAGAGCCGG.	GCAGTCTGAT

FIGURE 4

7/11

Osteocalcin Gene Sequence Comparison

	1				50
mouseCCA	ACCACAAGAA	ATGCCCTACA	ACCGGATCTT
rat	GCTCTCCCCA	TCAAAACCAC	TAAATAAGAA	ATGCCCTACA	ACCGGATCTT
human
	51				100
mouse	ATGGAGGCAT	TTTCTCAATT	GAGGTTTTCT	CCTTCCAAGA	AGACATAAAA
rat	ATGGAGGCAT	TTTCTCAATT	GAGGCTTTCT	TCTTTCAAGT	TGACAGAAAA
human
	101				150
mouse	CTAACCAGGA	CATCACCCCC	ACACACACAC	ACACACACCC	ACTGGATGAG
rat	CTAAGCAGGA	CACTTCCC..CAC	CCTTCTCCCC	ACTGGATGAG
human
	151				200
mouse	CAGAGCT.CC	CTGAACTGGG	CAAATGAGGA	CATTACTGAA	CACTCCCTCC
rat	CGGAGCTGCC	CTGCACTGGG	TGAATGAGGA	CATTACTGAC	CGCT.CCTTC
human
	201				250
mouse	CTGGGGTTTG	CGTCCCCTC	TCAGGGGCAG	ACACTGAAAA	TCACAGGCTA
rat	CTGGGGTTTG	GCTCCTGCTC	TCA.CGGCAG	ACACTGAAAA	TCACAGGCTA
humanACGGGGCTG	ACAGTAGAAA	TCACAGGCTG
	251				300
mouse	TGAG..AGTT	GGAGCCCAGG	TTATCCCAAA	CCGATTTTAG	ACTTCTGTAC
rat	TGCG..AGTT	GGAG.CCAGT	TTGTCCCAAA	CCGATTTTAG	ATCTCTGTAC
human	TGAGACAGCT	GGAGCCCAGC	TCTGCTTGAA	CCTATTTTAG	GTCTCTGATC
	301				350
mouse	CATGTCTAGG	C..ATGCATA	GGTTCTTGT	C.TCTAGGGC	GACCCAGTGC
rat	CCTCTCTAGG	CTATGCTCTA	AGGAGCTCGT	C.TCTAGGGC	CAGCCAGTGC
human	CCCGCTTCCT	CTTTAGACTC	CCCTAGAGCT	CAGCCAGTGC	TCAACCTGAG
	351				400
mouse	TCCAGCTGAG	GCTGAGAGAG	AGAGAGCACA	CAGTAGGAGT	GGTGGAGCAG
rat	TCCAGCTGAG	GCT..GAGAG	AGATGGCACA	CAGTAGGGGT	GCTGGAGCAG
human	GCTGGGGGTC	TCTGAGGAAG	AGTGAGTTGG	AGCTGAGGGG	TCTGGGGCTG
	401				450
mouse	CCCCTCAGGG	AAGAGGTCTG	GGGCCATGTC	AGAGCCTGGC	AGTCTCCGAT
rat	CCCCTCCGGG	AAGAGGTCTG	GGGCCATGTC	AGGACCCGGC	AGCCTCTGAT
human	TCCCCTGAGA	GAGGGGCCAG	AGGCAGTGTC	AAGAGCCGG.	GCAGTCTGAT

FIGURE 4

8/11

	451				500
mouse	TGTGGCCTCT	CGTCCA....	.CTCCCAGAG	CCTTGCCCAG	GCAGCTGCAA
rat	TGTGTCCTAC	CCTCCC....	.CTTCCAGGC	CTTCGCCCCG	GCAGCTGCAG
human	TGTGGCTCAC	CCTCCATCAC	TCCCAGGGGC	CCCTGGCCCA	GCAGCCGCAG
	501				550
mouse	TCACCAACCA	CAGCATCCTT	TGGGTTTGAC	CC.....A	CTGAGCACAT
rat	TCACCAACCA	CAGCATCCTT	TGGGTTTGAC	CT.....A	TTGC ^g CACAT
human	CTCCCAACCA	CAATATCCTC	TGGGGTTTGG	CCTACGGAGC	TGGGGCGGAT
	551				600
mouse	GACCCCAAT	TAGTCCTGGC	AGCATCCCCT	GCTCCTCCTG	CTTA.....
rat	GACCCCAAT	TAGTCCTGGC	AGCATCTCCT	GCCCCTCCTG	CTTA.....
human	GACCCCAAA	TAGCCCTGGC	AGATTCCCCC	TAGACCCGCC	CGCACCATGG
	601				650
mouseCAT	CAGAGAGCAC	AGAGTAGCCG	ATATAAATGC
ratCAT	..TAGGGGCT	CAGGCAGTGG	ATATAAAAGG
human	TCAGGCATGC	CCCTCCTCAT	CGCTGGGCAC	AGCCAGAGG	GTATAAACAG
	651				700
mouse	TACTGGATGC	TGGAGGGTGC	AGAACAGACA	AGTCCCACAC	AGCAGCTTGG
rat	TA.....TTGC	AGAACAGACA	AGTCCCACAC	AGCAACTCGG
human	TGCTGGAGGC	TGGCGGG.GC	AGGCCAGCTG	AGTCCTGAGC	AGCAGCCAG
	701				750
mouse	TGCACACCTA	GCAGACACCA	TGAGGACCCT	CTCTCTGCTC	ACTCTGTCTG
rat	TGCAGACCTA	GCAGACACCA	TGAGGACCCT	CTCTCTGCTC	ACTCTGTCTG
human	CGCAG.CCAC	CGAGACACCA	TGAGAGCCCT	CACACTCCTC	GCCCTATTGG
	751				800
mouse	CCCTGGCTGC	GCTCTGTCTC	TCTGACCTCA	CAGGTATGTG
rat	CCCTGACTGC	ATTCTGCCTC	TCTGACCTGG	CAGGTATGTG
human	CCCTGGCCGC	ACTTTGCATC	GCTGGCCAGG	CAGGTGAGTG	CCCCACCTC
	801				850
mouse	TCCTCCTGGT	TCATTTCTTT	GGG.....
rat	CCCTCCTGGT	TCATTTCTTT	GGA.....
human	CCCTCAGGCC	GCATTGCAGT	GGGGCTGAG	AGGAGGAAGC	ACCATGGCCC
	851				900
mouse	TAACCTA.CCT
rat	TAGCCACCCT
human	ACCTCTTCTC	ACCCCTTTGG	CTGGCAGTCC	CTTTGCAGTC	TAACCACCTT
	901				950
mouse	CCTGAAGG..TCTCA	CAATCTGCTT	TGGGATGGCA	GAGGGGAAGG
rat	CCTGAAGG..TCTCG	GGATCTGCTT	TGGGATGGCA	GAAGGGGAAGG
human	GTTGCAGGCT	CAATCCATTT	GCCCCAGCTC	TG ^g CCCTTGCA	GAGGGAGAGG

FIGURE 4, CONTINUED

9/11

	951				1000
mouse	GACA....AC	ACA.TGAGGG	AGACAGCAGG	GAGGAAACAG	AACTAAC...
rat	GACAAACTAC	CAG.AGAGAC	AGACAGCAGG	GAGGGAACAG	AACTAAC...
human	AGGGAAGAGC	AAGCTGCCCCG	AGACGCAGGG	GAAGGAGGAT	GAGGGCCCTG
	1001				1050
mouseTAC.CTGTT	TGCTTTACAG	ATGCCAAGCC
ratTGCACTGTT	TTCTTTACAG	GTGCAAAGCC
human	GGGATGAGCT	GGGGTGAACC	AGGCTCCCTT	TCCTTTGCAG	GTGCGAAGCC
	1051				1100
mouse	CAGCGGCCCT	GAGTCTGACA	AAGGTACTAG	CAGGAAGCCT	GGCAGGGCCT
rat	CAGCGACTCT	GAGTCTGACA	AAGGTACTGA	CAGGGAGCCT	GGCAGAACCT
human	CAGCGGTGCA	GAGTCCAGCA	AAGGTGCAGG	TATGAGGATG	GACCTGATGG
	1101				1150
mouse	CGGCTTGCC	TCAC.CCTGT	CCCCTAAGCC	CCCAAATCCC	CTTGCC....
rat	CAGCCTTGCC	CCTCACCCCTG	TCCCTAAGCC	CCCAGACCCC	TCTGCC....
human	GTTCTTGAC	CCTCCCTCT	CACCCTGGTC	CCTCAGTCTC	ATCCCCCAC
	1151				1200
mouseTTCTGCCTG	TGTCCCACTT
ratTTCTGTCTG	GGTGTCCAC
human	TCCTGCCACC	TCCTGTCTGG	CCATCAGGAA	GGCCAGCCTG	CTCCCCACCT
	1201				1250
mouse	TTCTCACTG	AACTCAGAAT	TACCTGAC.C	TTGTGTGTCT	TCTCCACAGC
rat	TTTC...CTC	CAAACCAGAA	TACCTGAT.C	CTGTGCCTC.	TGTTCCACAGC
human	GATCCTCCCA	AACCCAGAGC	CACCTGATGC	CTGCCCCCTCT	GCTCCACAGC
	1251				1300
mouse	CTTCATGTCC	AAGCAGGAGG	GCAATAAGGT	AGTGAACAGA	CTCCGGCGCT
rat	CTTCATGTCC	AAGCAGGAGG	GCAGTAAGGT	GGTGAATAGA	CTCCGGCGCT
human	CTTTGTGTCC	AAGCAGGAGG	GCAGCGAGGT	AGTGAAGAGA	CCCAGGCGCT
	1301				1350
mouse	AC.....CTTGGG	TAAG.....T	GCCAGAGCCC
rat	ACCTCAACAA	TGGACTTGGG	TGAGTGGGGA	GGAAGGCGCT	GCCAGGGCCC
human	ACCTGTATCA	ATGGCTGGGG	TGAGAGAAAA	GGCAGAGCTG	GGCCAAGGCC
	1351				1400
mouse	TTAGCCTTCC	ATATTGGTAG	GGAGGAGTTA	TGCTGGGGTG	GTTTCTGTGA
rat	TTTGCTTTCC	ATATTGGTAG	GGAGGAGTTG	TGCT.GGGTG	GTCTCTATGA
human	CTGCCTCTCC	GGGATGGTCT	GTGGGGGAGC	TGCAGCAGGG	AGTGGCCTCT
	1401				1450
mouse	CCCGCAGAGG	CTACACGTGC	AGGTCAATCC	CCATGTCCAG	GACCCTGGAG
rat	ACTGCCGAGG	GT.....C	CCATGCCTAG	GACCCTGGAG
human	CTGGGTGTG	GTGGGGGTAC	AGGCAGCCTG	CCCTGGTGGG	CACCCTGGAG

FIGURE 4, CONTINUED

10/11

	1451		1500
mouse	CCTCTTGTAC	AGTGTGGGAA	GAGGGTGTGT GTACCCCGTG TATATTAATG
rat	CCTCTTGTGC	A..GTGGGAA	GAGTGTGTGT ATACCCCGTG TTGGTTAATG
human	CCCCATGTGT	AGGGAGAGGA	G.GGATGGGC ATTTGTCACG GGGGCTGATG
	1501		1550
mouse	CCACTGTGTG	TTGGTTGATG	TTACTTTATA CTTCTCAGAG CCTCAGTCCC
rat	CCACTGCGTA	TTGGTTGACG	CTACTATATG CTTCTCAGAG CCCCAGCCCC
human	CCACCACGTC	GGG..... TGTCTCAGAG CCCCAGTCCC
	1551		1600
mouse	CAGCCCAGAT	CCCCTGGAGC	CCACCCGGGA GCAGTGTGAG CTTAACCCTT
rat	CTACCCAGAT	CCCCTGGAGC	CTCACAGGGA GGTGTGTGAG CTCAACCCC.
human	CTACCCGGAT	CCCCTGGAGC	CCAGGAGGGA GGTGTGTGAG CTCAA.TCCC
	1601		1650
mouse	GCTTGTGACG	AGCTATCAGA	CCAGTATGGC TTGAAGACCG CCTACAAACG
rat	AATTGTGACG	AGCTAGCGGA	CCACATTGGC TTCCAGGACG CCTACAAGCG
human	GACTGTGACG	AGTTGGCTGA	CCACATCGGC TTTCAGGAGG CCTATCGGCG
	1651		1700
mouse	CATCTACGGT	AT.CACTATT	TAGGACCTGT GCTGCCCTAA AGCCAAACTC
rat	CATCTATGGC	AC.CACCGTT	TAGGGCATGT GTTGCCCTGG AGCCCAACGC
human	CTTCTACGGC	CCGGTCTAGG	GTGTCGCTCT GCTGGCCTGG CCGGCAACCC
	1701		1750
mouse	TGGCAGCTCG	GCTTTGGCTG	CTCTCCGGGA CTTGATCCTC CCTGTCTCTT
rat	AG..CTTCAG	CTTTTGGCTA	CTCTCCAGGA CTCGACCCTC CCTGTTCCCT
human	CAG..TTCTG	CTCCTCTCCA	GGCACCCCTC TTCTCTCTTC CCCTTGCCCT
	1751		1800
mouse	CTCTCTGCCC	TGCAAG...T	ATGGATG... ..TCACAGCAG
rat	CTCTCTGCCT	CGAAAG...T	ATGGACG... ..GCACAGCTG
human	TGCCCTGACC	TCCCAGCCCT	ATGGATGTGG GGTCCCCATC ATCCCAGCTG
	1801		1850
mouse	CTCCAAAATA	AA.....
rat	CTCCAAAATA	AAGTCCAGAT	GAGGAACGGT TGGGCTCGAG TCTGTCCATT
human	CTCCAAAATA	AACTCCAGAA	GAGGAATCTG TGGGCCTGTG AGTCTGTCCA
	1851		1900
mouse
rat	GTGGTGTTGG	TAGGGGTTGG	AGAAGGCTTC A.TGTTTCATCC CAGACAGCCC
human	GTTTATGGAG	TGTGGGAGGG	AGGTGTCAGG A.GGATGGGGG TGAGGAGGTT
	1901		1950
mouse
rat	CAGTCCGGTT	ACTTTGCGTC	TTTATTATAC T.CTCTGTGA TGGATCGCAC
human	TTACCTTCTT	CAGTTCTAGA

FIGURE 4, CONTINUED

11/11

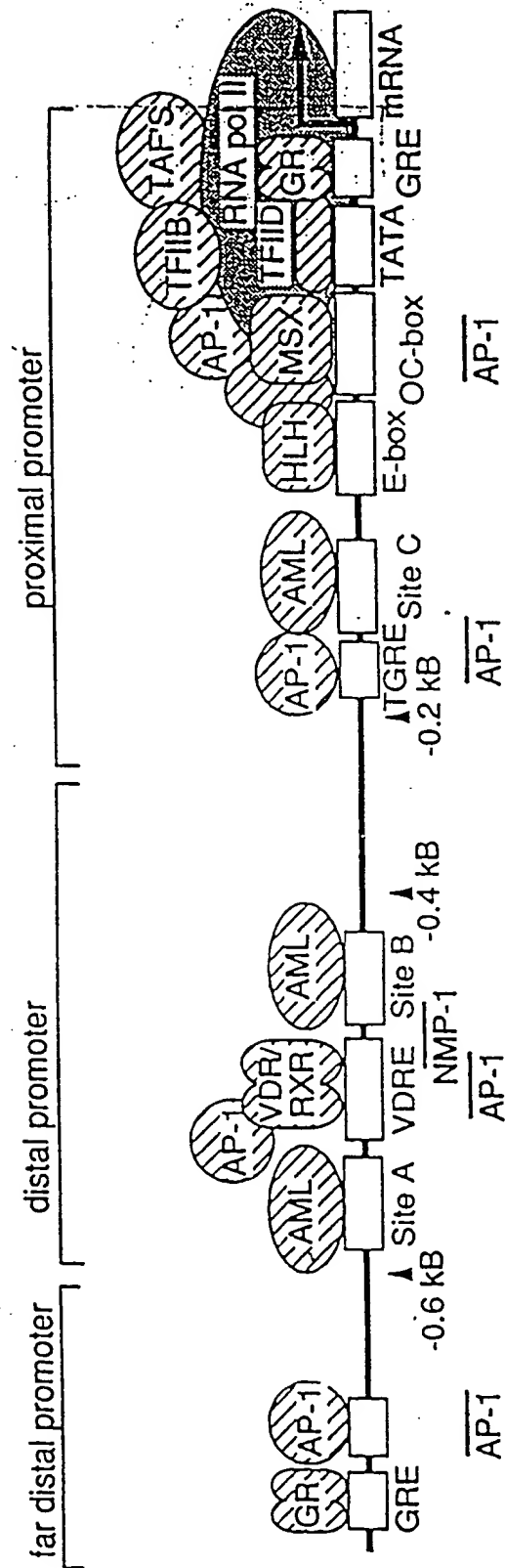


FIGURE 5